

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

Luiz Felipe de Lima Oliveira Silva

**PROPOSTA E AVALIAÇÃO DE SISTEMA DE
COMUNICAÇÃO DE INFORMAÇÃO ESCONDIDA EM
PROPAGANDAS SONORAS**

Florianópolis

2014

Luiz Felipe de Lima Oliveira Silva

**PROPOSTA E AVALIAÇÃO DE SISTEMA DE
COMUNICAÇÃO DE INFORMAÇÃO ESCONDIDA EM
PROPAGANDAS SONORAS**

Dissertação submetida ao Programa
de Pós-Graduação em Engenharia Elé-
trica para a obtenção do Grau de Mes-
tre em Engenharia Elétrica.

Orientador: Prof. Joceli Mayer, Ph.D.

Florianópolis

2014

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Silva, Luiz Felipe de Lima Oliveira

Proposta e avaliação de sistema de comunicação de
informação escondida em propagandas sonoras / Luiz Felipe
de Lima Oliveira Silva ; orientador, Joceli Mayer -
Florianópolis, SC, 2014.

109 p.

Dissertação (mestrado) - Universidade Federal de Santa
Catarina, Centro Tecnológico. Programa de Pós-Graduação em
Engenharia Elétrica.

Inclui referências

1. Engenharia Elétrica. 2. Marca d'água em áudio. 3.
Sistemas de comunicação. 4. Canal acústico. 5. Canal
ambiente. I. Mayer, Joceli. II. Universidade Federal de
Santa Catarina. Programa de Pós-Graduação em Engenharia
Elétrica. III. Título.

Luiz Felipe de Lima Oliveira Silva

**PROPOSTA E AVALIAÇÃO DE SISTEMA DE
COMUNICAÇÃO DE INFORMAÇÃO ESCONDIDA EM
PROPAGANDAS SONORAS**

Esta Dissertação foi julgada aprovada para a obtenção do Título de “Mestre em Engenharia Elétrica”, e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia Elétrica.

Florianópolis, 28 de março 2014.

Prof. Patrick Kuo Peng, Dr. Eng.
Coordenador

Banca Examinadora:

Prof. Joceli Mayer, Ph.D.
Orientador

Prof. Carlos Aurélio Faria Da Rocha, Dr.

Prof. Eduardo Luiz Ortiz Batista, Dr.

Prof. Rogério Richa, D.Sc.

A minha esposa Maynara, sem a qual minha vida seria incompleta

AGRADECIMENTOS

Agradeço ao professor Joceli Mayer, pelo apoio recebido e infraestrutura concedida. Em especial, pela orientação, pela paciência e pelas inúmeras sugestões, fundamentais para o desenvolvimento deste trabalho.

Agradeço ao professor Carlos Aurélio, de quem tive a honra de ser aluno na graduação, por sua participação na banca examinadora.

Agradeço ao professor Eduardo Luiz e à Rogério Richa, por sua participação na banca examinadora, e também a todos os amigos e colegas do LPDS.

Um agradecimento aos colegas da Certimarca, Samir Fiates, André Ribas, Rafael Scheidt, André Jobim e Gustavo Pereira. Especialmente ao Samir Fiates, pela ajuda, apoio e colaboração neste trabalho, e ao André Ribas, pelo apoio e suporte no desenvolvimento deste trabalho.

Agradeço a todos os meus familiares de Goiânia e de Florianópolis, em especial aos meus sogros Denise e Wilson, pelo apoio e auxílio quando necessitei, e aos meus tios e primos em Goiânia, por todo apoio e inspiração ao longo da minha vida.

Agradeço ainda, com extremo carinho, à minha esposa Maynara, pela sua paciência, apoio, ajuda e companheirismo, que sem dúvida me deram ainda mais motivação para a conclusão e auxiliaram na revisão deste trabalho.

Finalmente, gostaria de agradecer a minha mãe Marlene e à minha irmã Isa, pela ajuda emocional, carinho, convívio e auxílio, além da educação e apoio financeiro recebidos ao longo da minha vida.

Resumo

Esta dissertação propõe um novo sistema de comunicação informado projetado para incorporar a informação escondida em anúncios de áudio. A informação é modulada e transmitida através de uma marca d'água em áudio digital incorporada ao sinal de áudio hospedeiro. A inserção de marca d'água é projetada para alcançar um impacto perceptual mínimo e alta robustez às distorções do ambiente, a fim de cumprir com a proposta de aplicação em marketing. A energia da incorporação está devidamente dimensionada para atingir a alta transparência. Novas técnicas de sincronização e de equalização adaptativa são investigadas e propostas para extrair as informações de forma confiável após as interferências do canal ambiente. Os resultados experimentais, utilizando canais ambientes simulados e reais, ilustram o desempenho do sistema proposto e os resultados estão em contraste a outras técnicas existentes. O objetivo deste trabalho é expor um sistema de marca d'água digital de áudio que integre em um sistema de comunicação informada escondida em propagandas sonoras, permitindo a utilização deste sistema em aplicações de marketing.

Palavras-chave: marca d'água em áudio, sistemas de comunicação, canal acústico, canal ambiente.

Abstract

This dissertation proposes a new informed communication system designed to embed information hidden into audio advertisements. The information is modulated and transmitted through a audio digital watermark embedded into the audio signal host. The watermark embedding is designed to achieve minimal perceptual impact and high robustness to environment distortions in order to comply with the proposed application in marketing. The embedding energy is properly scaled in order to achieve high transparency. Novel synchronization and adaptive equalization techniques are investigated and proposed to reliably extract the information after the channel interferences. Experimental results using simulated and real environment channels illustrate the performance of the proposed system and results are contrasted to other existing techniques.

Keywords: audio watermarking, communication systems, acoustic channel, environment channel.

Lista de Figuras

Figura 1	Modelo padrão de um sistema de comunicação.....	36
Figura 2	Sistema de marcação interpretado em um modelo de comunicação com extrator cego.....	37
Figura 3	Sistema de marcação com sinal hospedeiro conhecido no extrator.....	37
Figura 4	Sistema de marcação como comunicação com informação extra no transmissor.....	39
Figura 5	Os sinais antes e depois do procedimento de codificação de fase.....	42
Figura 6	Um esquema típico de inserção do método de espalhamento em frequência.....	43
Figura 7	Limite absoluto audível no silêncio.....	45
Figura 8	Canais de transmissão de áudios digitais.....	47
Figura 9	Modelo do canal ambiente.....	52
Figura 10	Ondas refletem no ambiente e chegam atrasadas ao receptor.....	54
Figura 11	ISI na transmissão de pulsos.....	55
Figura 12	Modelo final do canal ambiente.....	56
Figura 13	Espectro de potência do sinal gravado e sinal estimado pelo modelo para diferentes canais.....	59
Figura 14	Esquema geral do sistema implementado.....	62
Figura 15	Diagrama do incorporador do sistema proposto.....	62
Figura 16	Esquema gráfico dos sinais incorporados ao sinal hospedeiro.....	64
Figura 17	Diagrama da busca da sincronização pelo sinal de sincronização.....	66
Figura 18	Diagrama do equalizador do sistema proposto.....	67
Figura 19	Diagrama do extrator do sistema proposto.....	69
Figura 20	Configuração básica de um filtro adaptativo.....	73
Figura 21	Diagrama de equalização de canal.....	79
Figura 22	Diagrama de identificação de canal inverso com filtragem adaptativa baseado na marca d'água.....	80
Figura 23	Comparação entre o uso ou não da codificação BCH...	84

Figura 24 Comportamento do sistema proposto frente canais simulados	86
Figura 25 Espectro de potência dos trechos dos áudios usados nos testes	88
Figura 26 Comportamento do sistema proposto em situação real .	90

Lista de Tabelas

Tabela 1	Características das publicações citadas	48
Tabela 2	Relação das oito posições relativas entre o emissor e o receptor do som.....	58
Tabela 3	BER média dos 8 canais.....	89

LISTA DE ABREVIATURAS E SIGLAS

DA	Digital-Analógico	25
AD	Analogico-Digital	25
HAS	Sistema Auditivo Humano (<i>Human Auditory System</i>) ..	25
PAM	Modelo Psicoacústico ISO-MPEG (<i>PsychoAcoustic Model</i>)	25
ODG	Gradiente de Diferença Objetiva (<i>Objective Difference Gradient</i>)	25
HESM	Modelo de Sensibilidade do Ouvido Humano (<i>Human Ear Sensitivity Model</i>)	25
ZFE	Equalizador Forçador a Zero (<i>Zero-Forcing Equalizer</i>) ..	26
RICE	Estimação de Canal com Interferência Reduzida (<i>Reduced Interference Channel Estimation</i>)	26
MMSEE	Equalizador por Minimização do Erro Quadrático Médio (<i>Minimum Mean Square Error Equalizer</i>)	26
bps	Bits por segundo (<i>bits per second</i>)	32
LSB	Codificação de <i>Bit</i> Menos Significativo (<i>Least Significant Bit</i>)	35
DFT	Transformada Discreta de Fourier (<i>Discret Fourier Transform</i>)	41
DCT	Transformada Discreta do Cosseno (<i>Discret Cosine Transform</i>)	43
DWT	Transformada Discreta Wavelet (<i>Discret Wavelet Transform</i>)	43
PRA	Matriz Bidimensional Pseudoaleatória (<i>Two-Dimensional Pseudo-Random Array</i>)	49
TSM	Modificação da Escala de Tempo (<i>Time Scale Modification</i>)	49
BER	Taxa de Erro de <i>Bit</i> (<i>Bit Error Rate</i>)	50
LCM	Mapeamento em Coordenadas Log (<i>Log Coordinate Mapping</i>)	51
ISI	Interferência Inter-simbólica (<i>Inter-Symbol Interference</i>) ..	54
PS	Espectro de potência (<i>Power Spectrum</i>)	57
PSD	Densidade Espectral de Potência (<i>Power Spectrum Density</i>)	68

LS	Método dos Mínimos Quadrados (<i>Method of Least Squares</i>)	71
MINT	Teorema Inverso de Múltiplas Entradas/Saídas (<i>Multiple-Input/Output Inverse Theorem</i>)	71
MMSEE	Equalizador por Minimização do Erro Quadrático Médio (<i>Minimum Mean Square Error Equalizer</i>)	71
DFE	Equalizador por Decisão Realimentada (<i>Decision-Feedback Equalizer</i>)	72
LSM	Mínimo Quadrado Médio (<i>Least Mean Squares</i>)	74
NLMS	Mínimo Quadrado Médio Normalizado (<i>Normalized Least Mean Squares</i>)	74
RLS	Mínimo Quadrado Recursivo (<i>Recursive Least Square</i>) ..	74

LISTA DE SÍMBOLOS

m	Mensagem.....	37
\hat{m}	Mensagem extraída.....	37
s	Sinal hospedeiro.....	37
w_a	Marca d'água.....	37
s_w	Sinal marcado.....	37
\hat{s}_w	Sinal marcado ruidoso.....	38
\hat{w}_a	Marca d'água ruidosa.....	38
c_a	Resposta em frequência do canal acústico.....	55
s_a	Áudio emitido.....	55
\hat{s}_a	Áudio resultante após passar pelo canal acústico.....	55
η_a	Ruído acústico.....	55
r	Sinal marcado recebido.....	56
c_{alto}	Resposta em frequência do canal dos alto-falantes.....	56
c_{mic}	Resposta em frequência do canal do microfone.....	56
c	Resposta ao impulso do canal ambiente.....	56
η_G	Ruído gaussiano branco.....	56
c_n	Resposta ao impulso do ruído ambiente.....	56
η_A	Ruído ambiente.....	56
p_c	Sequência pseudoaleatória do cabeçalho.....	62
p_b	Sequência pseudoaleatória dos <i>bits</i>	62
s_{sw}	Sinal hospedeiro com sinal de sincronização marcado.....	62
\bar{p}_c	Sequência pseudoaleatória do cabeçalho ajustada pelo HESM.....	64
\bar{p}_b	Sequência pseudoaleatória dos <i>bits</i> ajustada pelo HESM.....	64
α	Fator de compromisso entre robustez e transparência da marca d'água.....	64
w_s	Sinal de sincronização.....	65
r^s	Sinal recebido sincronizado.....	66
\bar{r}^s	Sinal sincronizado equalizado.....	67

Sumário

1 INTRODUÇÃO	23
1.1 TRABALHO PROPOSTO	25
2 APLICAÇÕES E PROPRIEDADES	29
2.1 APLICAÇÕES	29
2.1.1 Proteção de Propriedade	29
2.1.2 Comprovação de Propriedade Intelectual	29
2.1.3 Autenticação e Detecção de Adulteração	30
2.1.4 <i>Fingerprinting</i>	30
2.1.5 Controle de Cópias e de Acesso	31
2.1.6 Portador de Informação	31
2.2 PROPRIEDADES E CARACTERÍSTICAS	31
2.2.1 Transparência	32
2.2.2 Segurança	32
2.2.3 Capacidade de transmissão de informação	32
2.2.4 Robustez	32
2.2.5 Algoritmos de Detecção Cega ou Não Cega	33
2.2.6 Complexidade e Custo Computacional	33
2.2.7 Compromisso Entre Propriedades	33
2.3 CONCLUSÃO	34
3 MARCA D'ÁGUA DE ÁUDIO	35
3.1 MODELAGEM DA MARCA DE ÁUDIO	35
3.1.1 Modelo de Comunicação	36
3.1.1.1 Modelo Tradicional de Comunicação	36
3.1.1.2 Modelo Básico	37
3.1.1.3 Marca D'água como Comunicação Informada no Transmissor	38
3.2 ALGORITMOS DE MARCAÇÃO DE ÁUDIO DIGITAL	39
3.2.1 Codificação de <i>Bit</i> Menos Significativo (LSB)	39
3.2.2 Ocultamento de Eco	40
3.2.3 Codificação de Fase	41
3.2.4 Espalhamento Espectral	42
3.2.4.1 Propriedades do sistema auditivo humano (HAS)	44
3.2.4.1.1 <i>Mascaramento em Frequência</i>	44
3.2.5 <i>Patchwork</i>	45
3.3 MARCA D'ÁGUA DE ÁUDIO NO CANAL AMBIENTE	46
3.3.1 Levantamento Bibliográfico	47
3.3.2 Análise do Canal Ambiente	51

3.3.2.1 Canal Acústico	53
3.3.2.2 Canal Ambiente	55
3.4 CONCLUSÃO	57
4 SISTEMA PROPOSTO DE MARCA D'ÁGUA DE ÁUDIO	61
4.1 INCORPORADOR	62
4.1.1 Codificador e Modulador	63
4.1.2 Transparência	64
4.1.3 Sincronismo	65
4.2 SINCRONIZADOR	65
4.3 EQUALIZADOR	67
4.4 EXTRATOR	68
4.4.1 Demodulação	68
4.4.2 Decodificação	68
4.5 CONCLUSÃO	69
5 EQUALIZADOR ADAPTATIVO	71
5.1 FILTRAGEM ADAPTATIVA	73
5.1.1 Filtro de Wiener	75
5.1.2 LMS	76
5.1.3 NLMS	77
5.1.4 RLS	77
5.2 ESTIMATIVA DO CANAL AMBIENTE INVERSO	79
5.3 EQUALIZAÇÃO DO SINAL RECEBIDO	81
5.4 CONCLUSÃO	81
6 RESULTADOS	83
6.1 TESTE 1 - COMPARAÇÃO ENTRE O USO OU NÃO DO BCH	84
6.2 TESTE 2 - COMPORTAMENTO DO SISTEMA PROPOSTO EM CANAIS SIMULADOS	85
6.3 TESTE 3 - COMPORTAMENTO DO SISTEMA PROPOSTO EM SITUAÇÃO REAL (SALA DE ESCRITÓRIO)	89
6.4 CONCLUSÃO	91
7 CONCLUSÕES E PERSPECTIVAS	93
Referências	97

1 INTRODUÇÃO

Os avanços na tecnologia multimídia permitiram o surgimento de programas versáteis e de fácil usabilidade que, conjugados à queda nos preços de dispositivos digitais, garantiram a circulação de um volume imenso de informação na forma de arquivos de vídeo, áudio e imagens. Hoje o uso dos arquivos de vídeo, áudio e imagens é comum graças às ferramentas de compartilhamento e à *Internet*. Verifica-se uma mudança expressiva, tanto na economia mundial quanto no modo operandi de como a população se comunica, ocasionada por estes avanços nos meios de comunicação, nos computadores e nas tecnologias digitais.

São inúmeros os benefícios desta facilidade de divulgação e manipulação, bem como de troca e cópia de arquivos, como exemplos: a possibilidade de compartilhar vídeos, áudios e fotos com familiares, amigos ou conhecidos que estão distantes; a rápida troca de informações importantes; e um novo meio de propaganda atingir um alto quantitativo de pessoas que estão distantes.

Quanto ao último exemplo, cabe destacar que uma vez disponibilizadas na rede mundial de computadores, as propagandas podem ganhar vida própria, atingindo públicos e objetivos muitas vezes nem imaginados na sua criação. Além dos meios convencionais de difusão de publicidade (rádio, televisão, jornais, revistas, panfletos), são acessíveis novos canais de divulgação (redes sociais, sites de hospedagem de vídeos, *e-mail*) nos quais a comunicação entre companhias e usuários tornou-se mais dinâmica e interativa, permitindo o contato direto entre eles.

Ressalta-se que a divulgação baseada na interatividade do indivíduo, aliado ao aparecimento dos códigos de barra bidimensionais, tais como *DataMatrix* e *QRCode*, em jornais, revistas, panfletos e, até, propagandas na televisão demonstra que mudanças drásticas também afetam o mercado de preconício. Os códigos bidimensionais facilitaram uma proatividade que não é verificada nas propagandas convencionais, pois permitem acesso a mais informações ou, ainda, possibilitam uma dinâmica no relacionamento do indivíduo com o objeto, ou com o criador da publicidade. Os códigos bidimensionais facilitam essa proatividade ao direcionar o cliente a informações mais detalhadas, a conteúdos exclusivos ou a promoções.

Contudo, os códigos de barra têm desvantagens: ocupam um espaço que poderia ser utilizado para outros fins; afetam esteticamente

a publicidade; e são limitados quanto ao tipo de mídia (imagens e vídeos).

Em vista dessas desvantagens, esse tipo de publicidade se torna ainda mais atrativo com o uso de uma técnica que não afeta perceptualmente a mídia e é aplicável a um maior número de mídias, como áudio. Marca d'água digital é um exemplo desse tipo de técnica e cumpre esses requisitos. Dessa forma, durante a execução do preconício, o cliente pode, usando algum dispositivo digital (celular, *tablet*, computador pessoal), reconhecer, com o devido programa, a marca d'água presente. Então, será direcionado a algum tipo de conteúdo exclusivo dessa propaganda, despertando a curiosidade do indivíduo e a eficácia da divulgação para a empresa.

Embora a marca d'água tenha sido concebida com o objetivo de proteção da mídia digital (tanto de sua propriedade intelectual quanto de seu conteúdo), seu uso pode ser expandido para outras áreas, tais como o *marketing*. Isso envolve um processo de incorporação de mensagem digital perceptualmente transparente em um **sinal hospedeiro**, que carrega a mensagem de preconício com o objetivo de “inserir” a informação naquele sinal.

Nessa aplicação, o sinal digital é chamado de **marca d'água digital** e contém os dados que podem ser usados em inúmeras aplicações, incluindo gerenciamento de direitos digitais, monitoramento de difusão e prova de adulteração. Embora imperceptível, a existência da marca é indicada quando uma mídia marcada passa através de um extrator de marcas adequado.

Dentro da inserção de marcas d'água em áudio, existem sistemas específicos que apresentam melhor desempenho em situações específicas, tais como em áudios restritos ao ambiente digital, ou em áudios de um tipo de música específica ou em áudios gravados com ruído, etc. Um grande número de modificações em qualquer domínio pode ser usado como incorporação de marca d'água. As técnicas mais utilizadas fazem a inserção da marca d'água através da adição de ruído pseudoaleatório aditivo, o qual contém a informação desejada.

Em uma aplicação voltada ao *marketing*, a marca deve possuir alta transparência, a capacidade de transmissão de alguns *bits* por segundo e alta robustez a distorções do canal ambiente. Especificamente, essa robustez corresponde ao seguinte cenário: uma **fonte emissora** reproduz a propaganda, cujo áudio possui uma marca d'água transparente; o som percorre um caminho descrito por um **canal acústico**, no qual passa por um filtro e é somado ao ruído aditivo; chegando, por fim, ao **receptor** que grava o áudio e, possuindo o decodificador

adequado da marca d'água transparente, decodifica a marca e recupera o conteúdo inserido.

Nota-se que a fonte emissora é qualquer tipo de dispositivo que reproduza um som digital, como televisão, rádio e computador pessoal, realizando assim a conversão digital-analógico (DA) do áudio. Já o receptor é definido como todo aparelho capaz de realizar gravações do som ambiente, por exemplo celular, *tablet* e microfone conectado a um computador pessoal, executando a conversão analógico-digital (AD).

1.1 TRABALHO PROPOSTO

Neste trabalho será apresentado um sistema de inserção e extração de marcas d'água transparentes em áudios digitais e robusto aos efeitos do canal ambiente. Neste cenário existem três problemas:

1. garantir a transparência da marca d'água;
2. resistir aos efeitos do canal ambiente (filtragem e ruído aditivo);
3. e assegurar a sincronização do sinal recebido na extração da marca.

A primeira dificuldade possui muitas soluções na literatura, como o uso de modelos psicoacústicos do sistema auditivo humano (*Human Auditory System* - HAS), como o ISO-MPEG (*PsychoAcoustic Model* - PAM), para conformar a marca d'água de forma inaudível (CVEJIC, 2004; WOLFF; BARAS; SICLET, 2010); ou o ajuste da energia da marca segundo alguma métrica de transparência, por exemplo o gradiente de diferença objetiva (*Objective Difference Gradient* - ODG) (XIANG; HUANG, 2007). Neste trabalho, combinou-se ambas as técnicas, usando o modelo de sensibilidade do ouvido humano em função da frequência (*Human Ear Sensitivity Model* - HESM) (TERHARDT, 1979) para conformar as sequências pseudoaleatórias e o ajuste do ganho na adição dessas sequências ao sinal hospedeiro, de maneira a garantir um alto valor do ODG entre o sinal marcado e o original.

Quanto ao segundo obstáculo (resistir ao canal ambiente), o entendimento dos efeitos provocados pelo canal permite combatê-los de maneira eficaz. Assim, dada a natureza desse canal, sabe-se que as alterações são diversas e variáveis no tempo. No entanto, na literatura e neste trabalho, verificou-se uma aproximação aceitável dessas distorções por uma técnica mais simples, descrita por uma filtragem linear somada a um ruído colorido aditivo (STEINEBACH et al., 2002; WOLFF; BARAS; SICLET, 2010).

Com a simplificação dos efeitos do canal ambiente entendidos, em (WOLFF; BARAS; SICLET, 2010) foi proposto o uso de um equalizador forçador a zero (*Zero-Forcing Equalizer* - ZFE) usando o estimador de canal com interferência reduzida (*Reduced Interference Channel Estimation* - RICE) para fazer a estimativa do efeito da filtragem linear. Entretanto, esse estimador não trata de forma alguma o ruído colorido aditivo e necessita de um período prolongado de estimação do canal.

Diante disso, no sistema proposto, é utilizado o equalizador por minimização do erro quadrático médio (*Minimum Mean Square Error Equalizer* - MMSEE), realizando a estimativa direta do canal inverso com filtragem adaptativa, cujo sinal de treinamento é uma das sequências pseudoaleatórias da marca d'água, reduzindo-se o tempo de equalização. Esse equalizador possui maior resistência aos efeitos do ruído aditivo em comparação ao ZFE proposto em (WOLFF; BARAS; SICLET, 2010), sendo, portanto, um equalizador mais adequado para a situação apresentada.

Por último, a adversidade da correta sincronização do sinal recebido é contornada acrescentando um segundo sinal ao sinal hospedeiro. Esse sinal de sincronização é parte da marca d'água do sistema apresentado em (LIU; INOUE, 2003), que é robusto ao canal ambiente e transparente, porém possui baixíssima capacidade. Portanto, neste trabalho, esse sinal é usado somente para identificar o início da marca d'água incorporada.

Destaca-se que existem outras formas de sincronização na literatura, como a sincronização no tempo baseado no código de Bark (HUANG; WANG; SHI, 2002) e a extração de pontos de saliência (WU; SU; KUO, 2000). No entanto, a técnica proposta se adequou bem às exigências de apontar com precisão o início da marca d'água.

O Capítulo 2 introduz os conceitos básicos e definições de marca d'água digital, inicialmente apresentando aplicações e propriedades de diferentes métodos. A partir das características de cada técnica, definiu-se o método de marcação proposto neste trabalho.

No Capítulo 3, expõe-se a interpretação de marca d'água digital por um modelo de comunicação, explicando diferentes modelos e qual foi o utilizado neste trabalho. Ainda, é apresentado o estado da arte dos sistemas de marca d'água em áudio específicos ao canal analógico e ambiente ou robustos a estes. Dessa forma, é possível comparar o sistema proposto com o existente na literatura e determinar o modelo de canal ambiente aplicado neste trabalho.

Cada elemento do modelo proposto, o modelo e a comparação com os algoritmos originais são explicados no Capítulo 4. Tal capítulo

contêm diagramas explicando o sistema proposto e suas etapas de incorporação, sincronização, equalização e extração, assim como as explicações dos mesmos. Nesse capítulo, é possível perceber as contribuições do trabalho quando comparadas aos sistemas de estado da arte. Importante ressaltar que a etapa de equalização será abordada superficialmente nesse capítulo, cabendo no próximo tópico a sua abordagem detalhada.

Em seguida, o Capítulo 5 apresenta a equalização do sistema de estado da arte e como esta foi modificada para atender às necessidades do sistema proposto quanto ao período de estimação do canal e variabilidade do canal acústico. Além disso, as limitações quanto à estimativa do canal inverso usando filtragem adaptativa e os problemas da equalização por minimização do erro quadrático médio também são objetos de discussão.

O Capítulo 6 comenta os testes usados para demonstrar a robustez e transparência do método de marca d'água deste trabalho. Para tal, os resultados desses testes são indicados por gráficos, sendo realizados tanto para canais ambientes simulados quanto para canais ambientes reais.

Finalmente, o Capítulo 7 é destinado às conclusões referentes ao desenvolvimento e resultados deste trabalho, e quanto às perspectivas de surgimento de trabalhos futuros com o mesmo tema.

2 APLICAÇÕES E PROPRIEDADES

Marca d'água digital é considerada idealmente uma forma de comunicação de dados imperceptível, robusta e segura em relação ao sinal hospedeiro, incluindo inserção e extração deste. Basicamente, a informação da marca d'água incorporada permanece na mídia marcada e suporta modificações não intencionais, isto é, processos que afetam o sinal marcado, contudo, não visam eliminar a informação da marca. Igualmente, a marca pode ou não suportar tentativas intencionais de removê-la. O principal desafio no desenvolvimento de uma marcação é a inserção desta, de forma a detectá-la de maneira confiável. A importância das propriedades citadas depende significativamente da aplicação a qual o algoritmo é projetado.

Neste capítulo, algumas aplicações e também as principais propriedades para marca d'água digitais são comentadas (COX et al., 2008), observado que essas propriedades muitas vezes são conflitantes entre si.

2.1 APLICAÇÕES

2.1.1 Proteção de Propriedade

Em aplicações de proteção de propriedade, a marca contendo a informação de propriedade é incorporada no sinal hospedeiro. Espera-se que a marca, conhecida somente ao detentor dos direitos autorais, seja bastante robusta e segura (por exemplo, para sobreviver a processamentos comuns de sinal e ataques intencionais), permitindo ao detentor demonstrar a presença da marca no caso de uma disputa pela propriedade intelectual. Importante destacar que a detecção deve ter uma probabilidade de alarme falso muito pequena.

2.1.2 Comprovação de Propriedade Intelectual

É ainda mais importante utilizar a marca d'água para realmente prova de propriedade, do que usá-la somente na identificação da propriedade dos direitos autorais. O problema surge quando adversários usam programas de edição para substituir os direitos autorais originais pelos seus e, então, reclama o domínio do direito autoral para si. Para atingir o nível de segurança necessário para provar a propriedade, é

indispensável restringir o acesso ao detector.

Sem o detector, a remoção da marca pelo adversário é extremamente dificultada. Contudo, ainda que a marcação do dono não possa ser removida, um adversário pode tentar enfraquecer a capacidade de comprovar a propriedade intelectual. Para solucionar tal problemática, deve-se fazer uma modificação sutil: ao invés de se tentar provar diretamente a propriedade, por exemplo, ao inserir a assinatura "Luiz é dono desta imagem" na imagem hospedeira, deve-se inserir um algoritmo que busque provar que a imagem do adversário é derivada da imagem original marcada. Assim, é fornecida uma evidência indireta, sendo, por lógica, mais provável que o real proprietário possua a imagem em disputa, uma vez que da sua versão foram originadas as duas outras imagens.

2.1.3 Autenticação e Detecção de Adulteração

Nessa aplicação, insere-se um conjunto secundário de dados no sinal hospedeiro e, depois, emprega-se esse conjunto para determinar se o sinal hospedeiro foi ou não adulterado. A robustez contra a remoção da marca ou torná-la indetectável não é um problema, uma vez que não há motivação do ponto de vista do atacante. Por outro lado, falsificar uma marca de autenticação válida em um sinal hospedeiro não autorizado ou adulterado deve ser prevenido. Na prática, também é desejável localizar (na dimensão do tempo ou do espaço) e diferenciar a modificação não intencional de adulterações. Geralmente, a capacidade de transmissão de informação precisa ser alta para satisfazer o maior volume de dados em relação as aplicações de proteção de propriedade. A detecção é feita sem o sinal hospedeiro original (detecção cega), porque ou o original não é acessível ou sua integridade precisa ser estabelecida.

2.1.4 *Fingerprinting*

Aplicações de *fingerprinting*, ou impressão digital, usam informações adicionais incorporadas à marca para rastrear o autor ou os destinatários de uma cópia particular de arquivo multimídia. Os algoritmos desse tipo de aplicação precisam ter alta robustez a ataques intencionais e modificações por processamento de sinais como compressão com perdas ou filtragem. *Fingerprinting* também requer boas propriedades anti-conluio do algoritmo, por exemplo, não é possível inserir mais de

um número identificador no mesmo arquivo, caso contrário o detector não poderá discernir qual cópia está presente. A capacidade de transmissão necessária é similar à exigida nas aplicações de proteção de propriedade, isto é, de alguns *bits* por segundo.

2.1.5 Controle de Cópias e de Acesso

Em aplicações de controle de cópias, a marca d'água inserida representa um tipo de política de controle de cópia ou de acesso. O detector é, normalmente, integrado ao sistema de gravação ou reprodução. Após a detecção da marca e decodificação do conteúdo, a política de controle de cópia ou acesso é cumprida ao direcionar operações de *hardware* ou *software* específicas, como habilitar ou não o módulo de gravação. Essas aplicações necessitam de algoritmos resistentes a ataques intencionais e processamento de sinais, aptos para realizar detecção cega e capazes de incorporar um número incomum de *bits* no sinal hospedeiro.

2.1.6 Portador de Informação

O marcador nessa aplicação deve possuir alta capacidade e ser detectável e decodificável usando um algoritmo de detecção cega. Enquanto a robustez a ataques intencionais não é um requisito, um certo grau de robustez contra processamento comuns, como compressão, é desejado. Uma marca d'água pública inserida em um sinal hospedeiro pode ser usada para vincular informação adicional sobre o arquivo presente em um banco de dados externo, como informação de direitos autorais ou condições de licenciamento. Dados incorporados, por exemplo, em um vídeo de propaganda, podem conter informações sobre o produto ou direcionar a promoções exclusivas.

2.2 PROPRIEDADES E CARACTERÍSTICAS

As propriedades e características de marca d'água de áudio dependem da aplicação escolhida. Dessa forma, o propósito do sistema define o conjunto de requisitos da marca. Uma aplicação de propriedade intelectual exige que a marca resista a qualquer processamento aplicado. Por outro lado, a autenticação de dados impõe o oposto, que a marca seja perdida sob a aplicação de qualquer processamento. Assim,

seis aspectos principais para os algoritmos de marca d'água serão discutidos (COX et al., 2008; ARNOLD; SCHMUCKER; WOLTHUSEN, 2003).

2.2.1 Transparência

Na maioria das aplicações, o algoritmo de inserção da marca deve inserir dados adicionais sem afetar a qualidade perceptual do sinal hospedeiro. A transparência da marca em um sinal é maior, quanto maior for a fidelidade do sinal marcado ao sinal original, em outras palavras, quanto maior for a percepção de que não há diferença entre o sinal original e o marcado.

2.2.2 Segurança

A informação da marcação pode ter que ser sigilosa, somente pessoas autorizadas poderiam ser capazes de decodificar a marca. A segurança do processo de marcação é interpretada da mesma maneira que a segurança de criptografia é tratada. Nenhum indivíduo, mesmo que saiba que o sinal esteja carregando uma marca, deve ser capaz de extraí-la sem a chave de criptografia correta.

2.2.3 Capacidade de transmissão de informação

A capacidade de transmissão de informação, ou carga útil (*payload*), de uma marcação é o número de *bits* inseridos por uma unidade de tempo, geralmente dada em *bits* por segundo (*bits per second* - bps). A capacidade de transmissão da marca d'água é definido pelo tipo de aplicação e algoritmo utilizados na marcação.

2.2.4 Robustez

A robustez de um algoritmo é definida como a habilidade do detector da marca d'água de extrair a marca depois de manipulações ou ataques ao sinal marcado. Aplicações geralmente requerem robustez frente a um universo pré-definido de ataques. Por exemplo, em uma transmissão de rádio, uma marca deve sobreviver às distorções causados pelo processo de transmissão, incluindo compressão dinâmica e filtragem passa-baixa, pois a detecção é feita diretamente no sinal transmi-

tido. Por outro lado, em alguns algoritmos a robustez é completamente indesejada, são os chamados algoritmos de marcação frágeis, dentro da área de aplicação “Autenticação e Detecção de Adulteração” (ver Seção 2.1.3), que procuram identificar se o sinal teve algum tipo de processamento em seu caminho até o destinatário.

2.2.5 Algoritmos de Detecção Cega ou Não Cega

Em algumas aplicações, os algoritmos de detecção podem utilizar o áudio hospedeiro original para extrair a marca. Esta é a chamada detecção não cega. Porém se a aplicação não tem acesso ao sinal original, a extração deve ser feita somente com o sinal marcado (detecção cega).

2.2.6 Complexidade e Custo Computacional

A implementação de um sistema de marca d’água em áudio depende da atividade comercial envolvida na aplicação. O principal problema técnico para a complexidade computacional dos algoritmos de inserção e detecção de marcas é o requisito temporal. Enquanto em algumas aplicações a inserção e detecção deve ser feita em tempo real, em outras o tempo não é um fator crucial. Outro fator importante é o tipo de *software* e *hardware* na qual a implementação deve ser feita. É esperado que, com o crescimento dos dispositivos móveis, a demanda por aplicativos para essas plataformas aumente, porém há de se considerar o poder de processamento e memória desses aparelhos, o que muitas vezes limita o tipo de aplicação ou algoritmo utilizado.

2.2.7 Compromisso Entre Propriedades

É evidente a relação que existe entre as propriedades de um sistema de marca d’água. Por exemplo, pode-se aumentar a robustez de um sistema aumentando a energia de sua marca em relação ao sinal hospedeiro. Entretanto, isso compromete a transparência do sinal marcado frente o sinal original. Isto posto, esse exemplo demonstra que não faz sentido valorizar uma única propriedade esquecendo-se das demais. Logo, cada aplicação necessita trabalhar suas propriedades constituintes dando ênfase a uma ou outra, sem descuidar das demais, garantindo um compromisso entre o possível e o desejável, e entre as

características conflitantes de suas propriedades.

2.3 CONCLUSÃO

Neste capítulo foram descritas diversas aplicações que podem ser implementadas com o uso de marcas d'água. Foram também apresentadas as principais propriedades de sistemas de marca d'água, mostrando que as propriedades desejadas normalmente dependem do propósito, e que existe dependência entre elas.

O próximo capítulo apresenta os modelos de comunicação e como aplicá-los ao sistema de marca d'água. Ainda, especifica os algoritmos de marca d'água mais comuns usados em sistemas estado da arte, para finalizar com o estado da arte dos sistemas de marca d'água em áudio que ou são específicos ao canal analógico e ambiente ou são robustos a esses canais. Espera-se que o entendimento desses algoritmos sirva como fundamentação para compreensão do conteúdo diferenciado do Capítulo 4.

3 MARCA D'ÁGUA DE ÁUDIO

Nos capítulos anteriores foram discutidas as principais aplicações e propriedades, bem como um pouco da história das marcas d'água digitais. Em quase 30 anos desde o surgimento da marca d'água digital, diversas técnicas de marcação de áudio foram desenvolvidas e publicadas, variando desde seus objetivos de criação, métodos de elaboração, até sua aplicabilidade e forma de utilização. Sendo assim, hoje é possível interpretar essas marcas através de diferentes modelos, entre eles o **modelo de comunicação** ou o **modelo geométrico**, que serão examinados de maneira breve neste capítulo. Por outro lado, estas técnicas também podem ser agrupadas de outra forma, dividindo-as em 5 algoritmos diferentes (ARNOLD; SCHMUCKER; WOLTHUSEN, 2003; CVEJIC, 2004), os quais também são vistos neste capítulo. São eles: **codificação de bit menos significativo** (*Least Significant Bit - LSB*), **ocultamento de eco**, **codificação de fase**, **espalhamento espectral**, *patchwork*. Finalmente, trataremos a descrição do estado da arte em sistemas de marca d'água de áudio que trabalhem com canais analógico e/ou ambiente.

3.1 MODELAGEM DA MARCA DE ÁUDIO

O sistema de marcas d'água digitais possui diversas modelagens, isto é, interpretações de seu funcionamento, que podem ser divididas em dois grandes grupos: modelos baseados na interpretação da marcação como um método de comunicação e modelos baseados na interpretação geométrica dos algoritmos de marcação.

Esta seção abordará somente o modelo de comunicação visando uma breve contextualização do tema, relativo aos pontos de maior interesse desse trabalho. Para maiores esclarecimentos, recomenda-se a leitura de (COX et al., 2008) para encontrar uma descrição do modelo geométrico ou de (EGGERS; GIROD, 2002; CVEJIC, 2004) sobre o modelo alternativo de comunicação baseado na modelagem estatística da marca digital.

3.1.1 Modelo de Comunicação

Tendo em mente que um sistema de marcas d'água é uma forma de transmitir uma mensagem entre o incorporador da marca e o receptor da mesma, pode-se determinar que a marca d'água digital enquadra-se no **modelo tradicional de um sistema de comunicação** (COX et al., 2008). Sendo assim, primeiramente, trata-se a definição e ilustração desse modelo.

Em seguida, são revisados dois modelos, cuja diferença entre si reside em como o sinal hospedeiro é entendido dentro do modelo tradicional de comunicação. No **modelo básico**, o hospedeiro é interpretado puramente como ruído. No segundo modelo, o hospedeiro ainda é considerado um ruído, porém tal ruído é fornecido ao incorporador como **informação extra**, transformando o marcador básico no **marcador informado**. Há, ainda, mais um modelo de comunicação que considera o sinal hospedeiro não como um ruído, mas como uma segunda mensagem que precisa ser transmitida com a marca de forma multiplexada, porém este não será abordado, uma vez que não foi levado em conta para o desenvolvimento deste trabalho.

3.1.1.1 Modelo Tradicional de Comunicação

Os elementos básicos do modelo tradicional de comunicação são ilustrados na Figura 1. A mensagem de entrada \mathbf{m} , que se deseja transmitir através do canal de comunicação é codificada por um **codificador** no preparo para transmissão por meio do canal.

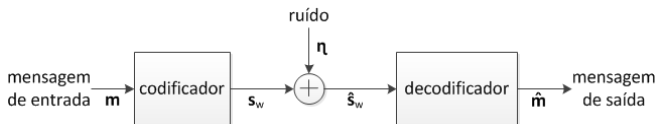


Figura 1 – Modelo padrão de um sistema de comunicação

A saída do codificador, \mathbf{s}_w , é enviada pelo canal de transmissão, assumido como ruidoso. Sendo assim, podemos considerar que o sinal recebido, aqui denominado $\hat{\mathbf{s}}_w$, é uma versão modificada do enviado, uma vez que o canal de transmissão agregou a característica de ruído ao sinal de saída. Como definição temos que essa mudança é entendida como a adição de um sinal ruidoso aleatório, $\boldsymbol{\eta}$, à \mathbf{s}_w .

Ao chegar no receptor, o sinal recebido entra no **decodificador**, este exercerá um processo inverso ao de codificação e, posteriormente, fará uma tentativa de correção dos erros de transmissão, resultando em uma mensagem de saída $\hat{\mathbf{m}}$, próxima ou igual a \mathbf{m} quanto menor for a probabilidade de erro do sistema.

3.1.1.2 Modelo Básico

As Figuras 2 e 3 apresentam uma forma de mapear a marca d'água seguindo o esquema da Figura 1. A Figura 2 mostra um sistema que usa um extrator cego, já a Figura 3 um sistema que usa um extrator informado (não-cego). Em ambos esquemas, a marcação é vista como um canal de transmissão no qual a mensagem é transmitida, considerando o sinal hospedeiro como parte desse canal.

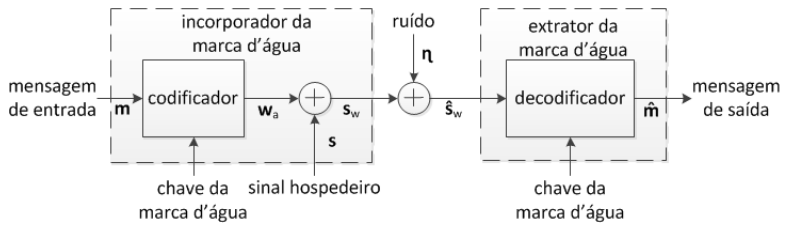


Figura 2 – Sistema de marcação interpretado em um modelo de comunicação com extrator cego

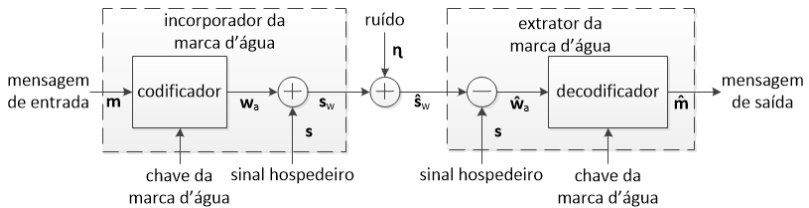


Figura 3 – Sistema de marcação com sinal hospedeiro conhecido no extrator

Nos dois mapeamentos, o processo de incorporação consiste em dois passos: primeiro, a mensagem \mathbf{m} é codificada em uma marca d'água \mathbf{w}_a , de mesmo tipo e dimensões que o hospedeiro \mathbf{s} . Essa codificação pode ser feita usando uma chave, a fim de torná-la mais segura.

Em seguida, \mathbf{w}_a é adicionado ao sinal hospedeiro \mathbf{s} , produzindo o sinal marcado \mathbf{s}_w . Esse marcador é chamado de **marcador cego**, pois o codificador ignora o hospedeiro, diferenciando-se do **marcador informado**, cujo codificador leva em consideração o sinal hospedeiro.

Após a incorporação do padrão, o sinal marcado \mathbf{s}_w sofre algum tipo de processamento, nesse caso, modelado como a adição de ruído. Os tipos de processamento que o sinal pode sofrer incluem compressão e descompressão, transmissão por canais analógicos, aprimoramento do áudio, entre outros. Entre esses processamentos, estão inclusas práticas maliciosas de terceiros com a intenção de remover a marca. Entretanto, para este trabalho essas práticas foram desconsideradas devido ao objetivo de aplicação da marca, como via de *marketing*, de maneira que tais procedimentos mostram-se irrelevantes.

Em um extrator cego, ilustrado na Figura 2, o sinal hospedeiro não-marcado é desconhecido, e, por isso, não pode ser removido antes da decodificação. Assim, analogamente à Figura 1, onde a mensagem codificada é corrompida pela adição de ruído, formado pela combinação do sinal hospedeiro e de um sinal ruidoso, o sinal marcado ruidoso $\hat{\mathbf{s}}_w$ é visto como uma versão corrompida da marca \mathbf{w}_a , e o extrator da marca é entendido como o decodificador.

No caso do extrator informado na Figura 3, o processo de detecção consiste em dois passos. Primeiro, o sinal hospedeiro é subtraído do sinal ruidoso $\hat{\mathbf{s}}_w$, para obter uma marca d'água ruidosa $\hat{\mathbf{w}}_a$. Esse padrão é decodificado pelo decodificador com a chave da marca d'água. Devido à adição do sinal hospedeiro no incorporador, sua interferência é precisamente cancelada pela sua subtração no extrator. Dessa forma, as diferenças entre \mathbf{w}_a e $\hat{\mathbf{w}}_a$ são causadas apenas pelo processo ruidoso. Sendo assim, pode-se ignorar a adição e a subtração do hospedeiro, o que significa que o codificador, o processo ruidoso e o decodificador juntos formam um sistema análogo ao sistema de comunicação da Figura 1.

3.1.1.3 Marca D'água como Comunicação Informada no Transmissor

Embora o modelo da Figura 2 nos contemple com alguns conceitos sobre a robustez da marca d'água com extratores cegos, esse modelo não abrange todos os algoritmos de incorporação possíveis, pois restringe o codificador como independente do sinal hospedeiro. Em consequência ao fato do sinal hospedeiro não-marcado \mathbf{s} ser visivelmente conhecido no marcador, não há razão para não usá-lo na marcação,

visto que há um aumento no número de possíveis técnicas de incorporação utilizáveis. Dito isso, a partir da Figura 4, visualizamos um modelo de marcação que permite que \mathbf{w}_a seja dependente de \mathbf{s} . Ainda que bastante semelhante ao modelo da Figura 2, notamos que a principal característica que os difere é o fornecimento de \mathbf{s}_w como uma entrada adicional no codificador.

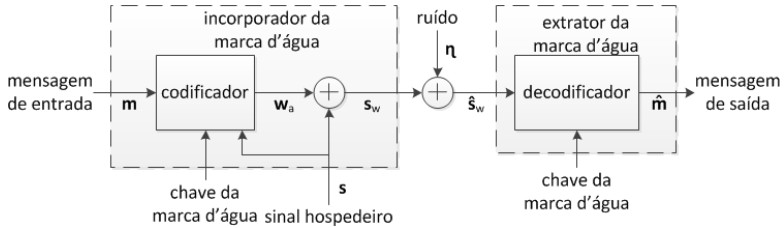


Figura 4 – Sistema de marcação como comunicação com informação extra no transmissor

Desde (SHANNON, 1958), alguns autores estudaram comunicação com informação extra (COSTA, 1983), inclusive para marca d'água digital (COX; MILLER; MCKELLIPS, 1999). Foi percebido que em alguns tipos de canais não há relevância na disponibilidade da informação extra, esteja ela disponível no transmissor (Figura 4), receptor (Figura 3) ou ambos, sua interferência pode ser eliminada. Por conseguinte, o sistema proposto usa o modelo da Figura 4, contudo a informação extra disponível no transmissor é utilizada para garantir a transparência da marca d'água.

3.2 ALGORITMOS DE MARCAÇÃO DE ÁUDIO DIGITAL

Há, na literatura, uma gama de algoritmos para incorporar informação no áudio digital. Esses algoritmos têm ampla abrangência, e entre eles existem técnicas como a técnica LSB e a técnica do espalhamento em frequência.

3.2.1 Codificação de *Bit* Menos Significativo (LSB)

Uma das primeiras técnicas a ser estudada na área do marca d'água de áudio, a LSB é particularmente simples em sua teoria e implementação (ARNOLD; SCHMUCKER; WOLTHUSEN, 2003). Tomando-se

uma definição a partir do próprio nome, essa abordagem utiliza os *bits* menos significativos do som para ocultar a informação. Já com a finalidade de aumentar a segurança da marcação, em muitas situações é gerada uma chave de criptografia que determina quais amostras serão selecionadas para a modificação dos *bits* menos significativos. Portanto, podemos determinar um subconjunto de amostras do áudio no qual são introduzidas modificações nos *bits* que traduzem a informação que se pretende inserir.

Sendo assim, a decodificação é feita de forma similar, através da chave secreta são selecionadas as amostras que foram modificadas e são lidos os *bits* menos significativos. A informação escondida no áudio é formada agrupando-se todos os *bits* e, uma forma de aumentar sua robustez, é repetindo a informação várias vezes em um mesmo sinal.

A modificação dos *bits* menos significativos introduz no sinal um ruído branco gaussiano aditivo de baixa intensidade. Devido à sensibilidade do sistema auditivo humano, o ruído introduzido pode torna-se perceptível com o aumento da quantidade de informação, limitando o número de *bits* que podem ser modificados imperceptivelmente (CVE-JIC, 2004).

A maior vantagem do método de Codificação de *Bit* Menos Significativo, é a alta capacidade de armazenagem de dados. Em um áudio amostrado a 44 kHz é possível armazenar 44 kbps usando somente um *bit*. Todavia, a desvantagem evidente é a baixíssima robustez da marca. Sendo assim, a marcação pode ser comprometida pela simples conversão DA/AD, pela adição de ruído ou ainda pela própria compressão da mesma.

3.2.2 Ocultamento de Eco

A técnica de ocultamento de eco é bastante explorada na literatura, principalmente pelas suas características que garantem uma alta qualidade do sinal marcado, o que torna a transparência da marca um ponto forte desse tipo de abordagem. Pode-se dizer que a diferença básica entre se ouvir um sinal de áudio sem marcação e um marcado com ocultamento de eco, é a mesma de se ouvir um som ao mesmo tempo com fones de ouvido e caixas de som. Com os fones só se ouve o som original, já com as caixas se ouve o som original somado com ecos causados pela acústica de um ambiente (GRUHL; BENDER; LU, 1996).

A abordagem original descrita em (GRUHL; BENDER; LU, 1996), utiliza somente um pulso de “eco” deslocado por um atraso δ , d_0 ou

d_1 dependendo do *bit* que se quer inserir, e multiplicado por um fator α de decaimento. Portanto, o áudio é dividido em várias partes e em cada parte é adicionado um “eco” com um atraso que representa o *bit* inserido. Na detecção o sinal também é dividido e a identificação dos “ecos” se dá pela análise do cepstro de cada parte do sinal (GRUHL; BENDER; LU, 1996). É verificado um grande pico no gráfico do cepstro exatamente sobre o atraso que foi inserido durante a marcação, se o pico está em d_0 , o *bit* codificado é 0, caso esteja em d_1 , o *bit* é 1. Assim, após verificar todas as partes do áudio, pode-se obter a informação que foi inserida pelo processo de marcação de áudio digital.

Apesar da vasta literatura trazendo melhorias para essa técnica, verifica-se que a robustez da marcação ainda é uma das principais desvantagens se comparada a abordagens que utilizam o domínio da frequência para ocultar informação. Isto acontece uma vez que a informação está inserida no domínio do tempo, sendo facilmente corrompida pela compressão, adição de ruído, entre outros tipos de processamentos muito comuns no mundo do áudio digital.

3.2.3 Codificação de Fase

Os métodos de codificação por fase se aproveitam da insensibilidade do sistema de audição humano às mudanças absolutas na fase de um sinal para transmitir a informação. A técnica atua substituindo a fase do início de um segmento por uma de duas fases de referência. Estas fases de referência estão relacionadas com os *bits* 0 e 1. Para manter a distorção de fase no menor nível possível, são feitos ajustes nas fases dos segmentos adjacentes para que a diferença relativa entre eles seja mantida. A inserção é feita da seguinte forma:

1. O sinal original é dividido em pequenos segmentos cujo o tamanho é igual ao da mensagem que se quer inserir;
2. A transformada discreta de Fourier (*Discret Fourier Transform - DFT*) é aplicada em cada segmento e é obtida uma matriz com magnitude e fase;
3. A diferença entre fases adjacentes é calculada;
4. A mudança de fase entre segmentos consecutivos é facilmente detectável. Em outras palavras, o valor absoluto da fase dos segmentos pode ser mudada, porém a diferença relativa entre fases de segmentos adjacentes deve ser preservada. Portanto a informação

só é inserida no vetor de fase do primeiro segmento do sinal. Coloca-se uma fase de 90° para um *bit* 1 e -90° para um *bit* 0, por exemplo;

5. Uma nova matriz é criada usando a nova fase do primeiro segmento e as diferenças originais entre as fases;
6. Usando a nova matriz de fase e a matriz original da magnitude, o sinal é reconstruído aplicando a DFT inversa e concatenando os segmentos.

Para fazer a extração, o decodificador deve estar informado do tamanho exato do segmento. Então é feita a DFT para se conseguir a fase do sinal e assim retirar a informação. No entanto, uma grande desvantagem deste método é a baixa capacidade do algoritmo, visto que a mensagem é inserida somente no primeiro segmento do sinal (DUTTA; BHATTACHARYYA; KIM, 2009).

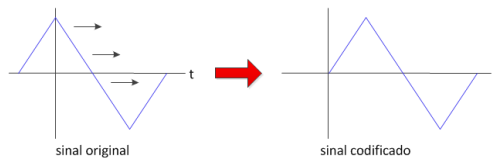


Figura 5 – Os sinais antes e depois do procedimento de codificação de fase

3.2.4 Espalhamento Espectral

O espalhamento em frequência é um exemplo de método de correlação que insere uma sequência pseudoaleatória e detecta a marca d'água calculando a correlação entre a sequência de ruído pseudoaleatório e o áudio marcado. Contudo, ainda que seja um método bastante simples para implementação, suas desvantagens são significativas, uma vez que há um grande custo computacional para criar um modelo psicoacústico para reduzir o ruído audível, e é suscetível a ataques de modificação da escala temporal.

De forma sucinta, podemos definir o conceito desse método como uma forma de espalhar uma sequência pseudoaleatória por todo o sinal de áudio. O ruído de banda larga pode ser espalhado tanto no domínio do tempo quanto no domínio transformado, não importando

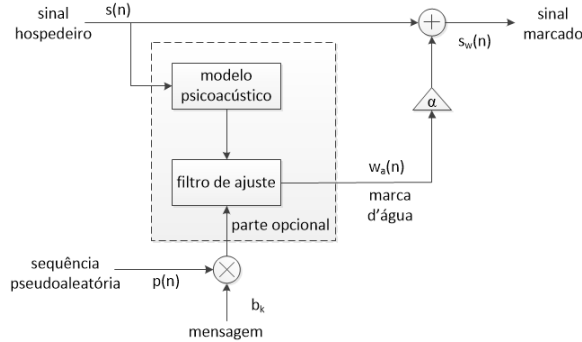


Figura 6 – Um esquema típico de inserção do método de espalhamento em frequência

que transformada seja usada. Transformadas usadas frequentemente são a transformada discreta do cosseno (*Discret Cosine Transform* - DCT), transformada discreta de Fourier (DFT) e transformada discreta *wavelet* (*Discret Wavelet Transform* - DWT). A mensagem binária da marca d'água $v = \{0, 1\}$ ou o equivalente bipolar $b = \{-1, +1\}$ é modulado por uma sequência pseudoaleatória $p(n)$ gerada por uma chave secreta.

Então, a marca modulada $w_a(n) = b_k p(n)$, sendo $p(n)$ uma sequência pseudoaleatória, é escalada pelo fator α baseado no compromisso entre robustez e transparência da marca, com a opção de ser conformada segundo o modelo psicoacústico do sinal de áudio. Conforme evidenciado na Figura 6, o sinal modulado é somado ao áudio original para produzir o áudio marcado $s_w(n)$. Dessa forma,

$$s_w(n) = s(n) + \alpha w_a(n). \quad (3.1)$$

A detecção dos *bits* inseridos usa a abordagem da correlação linear, dado que a sequência pseudoaleatória $p(n)$ é conhecida e pode ser regenerada através da chave secreta. As marcas são então detectadas usando a correlação entre $s_w(n)$ e $p(n)$

$$\begin{aligned} \rho &= \frac{1}{N} \sum_{i=1}^N s_w(i) p(i) \\ &= \frac{1}{N} \sum_{i=1}^N s(i) p(i) + \frac{1}{N} \sum_{i=1}^N \alpha b_k p^2(i), \end{aligned} \quad (3.2)$$

sendo N o comprimento do sinal. Assume-se que o primeiro termo da Eq. (3.2) é, com grande probabilidade, muito pequeno, uma vez que $s(n)$ e $p(n)$ são descorrelacionados. Entretanto, nem sempre são descorrelacionados. Se o primeiro termo tiver magnitude similar ou maior que o segundo termo, a detecção será incorreta. Usando o valor da correlação ρ e um limite pré-definido τ , as saídas detectadas são

$$\hat{b}_k = \begin{cases} 1 & \text{se } \rho > \tau \\ 0 & \text{se } \rho \leq \tau \end{cases}$$

Ao contrário do ocultamento de eco, o espalhamento espectral é bastante robusto ao ruído aditivo, a alguns tipos de compressão, a certos processamentos que o sinal pode sofrer (dependendo do método de espalhamento utilizado), dentre outros. Por outro lado, a aplicação da parte opcional da Figura 6 garante um grau de transparência equivalente ao ocultamento de eco ao custo do aumento computacional deste método.

3.2.4.1 Propriedades do sistema auditivo humano (HAS)

A percepção auditiva é baseada na análise de banda crítica no ouvido interno, onde a transformação da frequência em localização acontece ao longo da membrana basilar (CVEJIC, 2004). O espectro de potência dos sons recebidos não são representados por uma escala linear de frequência, mas por limitadas faixas chamadas **bandas críticas** (FASTL; ZWICKER, 2007). O sistema auditivo humano é normalmente modelado como um banco de filtros passa-faixa, consistindo em filtros fortemente sobrepostos com larguras de banda por volta de 100 Hz por faixa para frequências centrais abaixo de 500 Hz e com larguras de até 5 kHz por faixa para bandas em alta frequências. Sendo assim, se a maior frequência é limitada a 24 kHz, 26 bandas críticas devem ser consideradas.

3.2.4.1.1 Mascaramento em Frequência

Uma propriedade do HAS usada nos algoritmos de marca d'água é o **mascaramento em frequência** (FASTL; ZWICKER, 2007), definido como um fenômeno no domínio da frequência no qual um sinal de baixa intensidade, como uma **melodia pura (o mascarado)**, pode se tornar inaudível (mascarado) por um sinal mais forte simultâneo (**o**

mascarador), como um **ruído de banda estreita**, se o mascarador e o mascarado são próximos o suficiente entre si em frequência (FASTL; ZWICKER, 2007). Dessa maneira, um limite de mascaramento pode ser obtido abaixo do qual qualquer sinal não será audível, sendo que este depende das características do mascarador e do mascarado (ruído de banda estreita ou melodia pura).

Um modelo HAS largamente utilizado em marca d'água de áudio é **modelo psicoacústico ISO-MPEG** (PAM), baseado em fundamentos psicoacústicos e entropia perceptual (PAINTER; SPANIAS, 2000; COX et al., 2008). Esse modelo associa o limite de mascaramento citado ao modelo de sensibilidade do ouvido humano (HESM) em função da frequência (TERHARDT, 1979). O gráfico da Figura 7 mostra o nível mínimo de pressão sonora audível, o qual é recíproco à sensibilidade, para cada frequência.

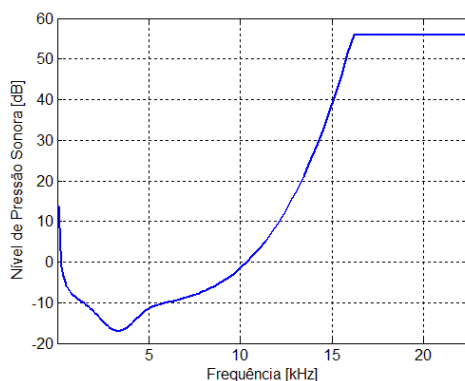


Figura 7 – Resposta em frequência do ouvido humano baseado no modelo de Terhardt (TERHARDT, 1979). Essa curva representa o limite absoluto, ou o limite de silêncio, para uma média de jovens ouvintes com audição apurada

3.2.5 *Patchwork*

A técnica de *patchwork* utiliza métodos estatísticos baseados em testes de hipóteses para transmitir a informação oculta, sendo, originalmente, voltada para a marca d'água de imagens. O método de inserção da marca utiliza um processo pseudoaleatório para inserir uma deter-

minada estatística dentro do áudio hospedeiro. A marca é extraída com a ajuda de indexadores numéricos, como a média e variância, que descrevem a distribuição estatística. É normalmente aplicada no domínio transformado (Fourier, *Wavelet*, entre outros) para espalhar a marca no domínio do tempo e aumentar a robustez contra ataques e modificações (CVEJIC, 2004). Sendo assim, o algoritmo apresenta resultados satisfatórios em relação a audibilidade do som marcado e a robustez em geral, porém encontra sérios problemas frente a ataques de dessincronização (KIM et al., 2004).

3.3 MARCA D'ÁGUA DE ÁUDIO NO CANAL AMBIENTE

A marca d'água digital de áudio desse trabalho é interpretada como um modelo de comunicação, especificamente, como o modelo de comunicação informado no transmissor. Em outras palavras, a marca d'água é inserida em um incorporador informado e extraída em um detector cego, sendo modificada entre a saída do primeiro e a entrada do segundo. Esta modificação equivale ao canal de transmissão da comunicação que pode ser descrito para o som digital como os cenários apresentados em (BENDER et al., 1996), como mostrado na Figura 8.

O primeiro sinal é transmitido através do canal de maneira inalterada, segundo a Figura 8(a). Isto posto, nenhuma característica do sinal transmitido é modificada. Na Figura 8(b), o sinal transmitido é re-amostrado resultando em um sinal recebido com uma taxa de amostragem diferente do sinal original. Isso significa que a amplitude e a fase permanecem inalteradas, mas as características temporais mudam. O terceiro caso, na Figura 8(c), converte o sinal e o transmite na forma analógica. Nessa situação, mesmo se o canal é considerado razoavelmente limpo, a amplitude, a fase e a taxa de amostragem do sinal transmitido são modificados. Na última situação, vista na Figura 8(d), o canal é semelhante ao anterior, exceto por não ser limpo. O sinal é sujeito a transformações diversas, inclusive não-lineares, resultando em mudanças de fase, mudanças de amplitude, eco, reverberação, ruído, entre outras alterações.

Em suma, na Figura 8(a), o áudio marcado não é afetado pelo **canal digital**; já na Figura 8(b), o áudio é re-amostrado ao passar pelo **canal re-amostrado**; a transformação DA/AD, com ou sem ruído, que aparece na Figura 8(c), é o chamado **canal analógico** e afeta áudio marcado de diversas maneiras; finalmente, o pior caso acontece quando o áudio é atacado por várias interferências simultaneamente ao

sofrer a transformação DA/AD e ser transmitido pelo ar livre, chamado **canal acústico**, exemplificado na Figura 8(d) e denominado **canal ambiente**.

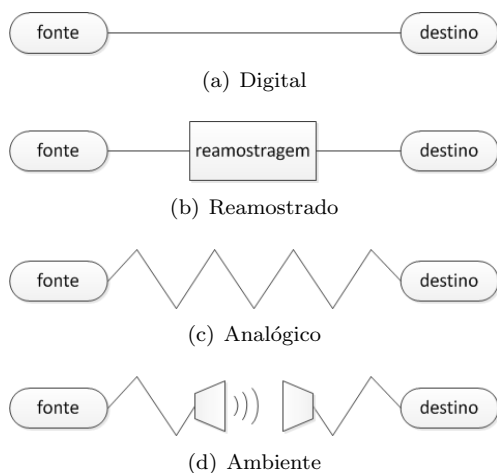


Figura 8 – Canais de transmissão de áudios digitais

As publicações na área de marca d'água digital de áudio focam-se principalmente nos efeitos do primeiro e segundo canais de transmissão. Contudo, alguns trabalhos consideram o terceiro canal, e ainda menos publicações o quarto cenário. Em muitos trabalhos, esses últimos dois cenários são citados apenas brevemente ressaltando sua robustez (HAITSMA et al., 2000; NAKAYAMA et al., 2003; LIE; CHANG, 2006; KANG; YANG; HUANG, 2011). Porém, muitos artigos consideram aplicações nas quais é necessária a sobrevivência da marca ao canal analógico ou ao canal ambiente (STEINEBACH et al., 2002; SHIN et al., 2002; TACHIBANA, 2003; LIU; INOUE, 2003; XIANG; HUANG, 2006; HUANG et al., 2007; WOLFF; BARAS; SICLET, 2010; QIAN et al., 2010; XIANG et al., 2011). Essas características estão listadas na Tabela 1 e, em seguida, os trabalhos citados são comentados brevemente e conclusões sobre as propriedades do sistema de marca d'água proposto são listadas.

3.3.1 Levantamento Bibliográfico

Um algoritmo de espalhamento em frequência com ponderação no tempo voltado para monitoramento e controle de cópias foi apresen-

Tabela 1 – Características das publicações citadas

Publicação	Objetivo		Canal analisado	
	Sobre-viver	Outro	Analógico	Ambiente
(HAITSMA et al., 2000)		X		X
(STEINEBACH et al., 2002)	X			X
(SHIN et al., 2002)	X		X	
(NAKAYAMA et al., 2003)		X	X	
(TACHIBANA, 2003)	X			X
(LIU; INOUE, 2003)	X		X	
(LIE; CHANG, 2006)		X	X	
(XIANG; HUANG, 2006)	X		X	
(HUANG et al., 2007)	X			X
(WOLFF; BARAS; SICLET, 2010)	X			X
(QIAN et al., 2010)	X			X
(KANG; YANG; HUANG, 2011)		X		X
(XIANG et al., 2011)	X		X	

tado em (HAITSMA et al., 2000). No trabalho, afirma-se que a marca d'água é robusta, resistindo a vários processamentos, inclusive o canal ambiente (reprodução por alto-falantes, DA; transmissão pelo canal acústico; gravação por microfone, AD). Quando em 2002, surge uma das primeiras referências que busca a caracterização do cenário da Figura 8(d) e procura criar um modelo para testes automatizados de algoritmos de marcas d'água com o programa Stirmark Benchmark (STEINEBACH et al., 2002). Concluiu-se que o canal ambiente causa, de forma geral, distorções de quantização, adição de ruído, ganho ou atenuação do sinal, adição de sinais externos (ruído de fundo, conversa, sons em geral) e filtragem no áudio transmitido. Entretanto, o estudo não deixa claro qual o método de marcação adotado, limitando a importância deste trabalho.

Ainda em 2002, é proposta outra abordagem em (SHIN et al., 2002), a qual utiliza a técnica que extrai do sinal de áudio uma portadora da qual a marca, que contém informação de direitos autorais, é produzida. Essa portadora é o resultado da filtragem passa-baixas do áudio, e a marca é gerada a partir da mudança do conteúdo da frequência utilizando a operação na escala de tons (*pitch scaling*), tornando-se resistente contra conversão DA/AD e compressão com perda.

No ano seguinte, (TACHIBANA, 2003) monitora shows ao vivo gravados (canal ambiente) usando uma marca d'água oculta, que emprega o algoritmo de espalhamento em frequência através de uma matriz bidimensional pseudoaleatória (*Two-Dimensional Pseudo-Random Array - PRA*) (TACHIBANA et al., 2002). (TACHIBANA et al., 2002) propôs uma técnica robusta a deslocamento e oscilações no tempo e na frequência (*pitch shifting* e *wow-and-flutter*) e corte aleatório de amostras. Essa robustez é atingida principalmente através da PRA, da modificação da magnitude e das subfaixas não-lineares. Posteriormente, notou-se sua sobrevivência à compressão MPEG, ruído aditivo, eco e conversão DA/AD.

(NAKAYAMA et al., 2003), assim como (HAITSMA et al., 2000), incorpora a marca no domínio da frequência, baseado-se no método de (BONEY; TEWFIK; HAMDY, 1996). Porém os resultados são referentes ao canal analógico (conversão DA/AD e transmissão por um cabo coaxial); e o algoritmo possui maior transparência, com o uso da técnica de mascaramento sucessivo com base no modelo psicoacústico MPEG, e melhor detecção, baseada o método de correlação cruzada branqueada (OMOLOGO; SVAIZER, 1996). Propondo uma nova maneira de utilizar padrões sinusoidais para criar uma marca d'água de espalhamento de frequência, (LIU; INOUE, 2003) explica como usar sequências pseudoaleatórias para gerar esses padrões sinusoidais de maneira que tenham as mesmas propriedades correlacionais das sequências pseudoaleatórias, e demonstra a robustez desta técnica a inúmeros ataques, inclusive conversão DA/AD (canal analógico), e sua alta transparência (teste ABX).

Em 2006, dois importantes autores (LIE; CHANG, 2006) propuseram que o sinal de áudio fosse modificado no domínio do tempo baseado no princípio de diferencial de amplitudes. Embora seja temporal, há um teste de mascaramento em frequência para garantir a inaudibilidade da marcação. Os autores demonstram que além de resistente ao canal analógico, também resiste à compressão com perda, normalização de amplitude, filtragem passa-baixas, modificação da escala de tempo (*Time Scale Modification - TSM*) e reamostragem. Assim como (STEINEBACH et al., 2002), (XIANG; HUANG, 2006) analisaram o canal analógico (conversão DA/AD e transmissão por um cabo) frente a marca de áudio baseada em quantização. Embora tenham concluído que essa marca é frágil frente a tal canal, caracterizou-se alguns dos efeitos deste canal: degradação por ruído, transformação temporal e modificação de amplitude.

Já para (HUANG et al., 2007), a marca incorporada possui o propósito de fornecer informações aos passageiros em um carro por

meio do sinal de áudio do rádio do carro. Desse modo, o algoritmo deve ser robusto ao canal ambiente e ao canal analógico AM/FM, levando os autores a utilizar a DCT para adicionar a marca. O principal aspecto desse método é sua alta taxa de transmissão, atingindo 430 bps. Entretanto, o principal déficit do artigo foi a falta de um teste real para a transmissão AM/FM. O algoritmo aplicado por (HUANG et al., 2007) baseia-se no estudo de (LAZIC; AARABI, 2006), cujas mudanças referem-se ao uso da DCT e de uma sequência pseudoaleatória com elementos uniformemente distribuídos no intervalo $[0,1]$ ao invés da DFT e de uma sequência pseudoaleatória binária.

Contrariando os trabalhos já citados, outro autor nos propõe incorporar a marcação no domínio do tempo, aplicando o modelo psicoacústico para atingir transparência, e usando o filtro não-causal de Wiener no detector [algoritmo encontrado em (BARAS; MOREAU; DYMARSKI, 2006)] e estimação do canal acústico (WOLFF; BARAS; SICLET, 2010). A sua maior contribuição é justamente o uso da estimação de canal com interferência reduzida (RICE) para a equalização do sinal (anulação dos efeitos do canal) antes da extração da marca. Porém os autores consideram que o sinal recebido está perfeitamente sincronizado desde o início, o que não é sempre possível na prática. Almejando um método de marcação robusto ao canal ambiente, transparente e de baixa complexidade computacional, (QIAN et al., 2010) propuseram alterar as magnitudes da DFT para inserir a marca, tomando o cuidado de maximizar a energia das magnitudes sem comprometer a máscara do PAM. Segundo os autores, o sincronismo da DFT empregado não é satisfatório e a transparência poderia ser melhorada, todavia os resultados respeitaram as condições necessárias para a aplicação.

Dando continuidade a sua pesquisa de 2006, (XIANG et al., 2011) nos propõe um algoritmo baseado na DWT para combater o canal analógico analisado em (XIANG; HUANG, 2006). O método adota estratégias de marcação das baixas frequências da DWT, alteração de forma relativa das energias nestas subfaixas e adição de códigos de sincronismo e aplicação de uma operação de interpolação para combater, respectivamente, o ruído aditivo degradante, as mudanças da energia do sinal de áudio na conversão DA/AD e modificação da escala de tempo. Na comparação feita no trabalho, a técnica proposta consegue uma melhora significativa na BER para diversas situações, inclusive o canal analógico, frente a outros 6 algoritmos. Já em 2011, outro trabalho mostrou taxas de erro de *bits* (*Bit Error Rate* - BER) nulas ou até 2% para o canal ambiente e, mais uma vez, adicionou a marca d'água no domínio da frequência utilizando o recurso geométrico inva-

riante de mapeamento em coordenadas log (*Log Coordinate Mapping* - LCM) (KANG; YANG; HUANG, 2011). Esse estudo evidencia que o recurso LCM sobrevive a cortes aleatórios do áudio em até 95%, deslocamento de tom (*pitch shifting*) de $\pm 20\%$, TSM em até $\pm 20\%$ e canal ambiente. Ainda, provou-se que esse algoritmo é robusto à filtragem passa baixa, compressão MP3, adição de eco, mudanças de amplitude e normalização.

Frente ao encontrado na literatura e aos estudos referentes ao canal ambiente discutidos anteriormente, pode-se concluir, no que se refere às propriedades do sistema proposto, quatro principais características:

- usar algum modelo psicoacústico do HAS;
- utilizar espalhamento em frequência;
- reduzir as modificações causadas pelo canal ambiente;
- e aplicar códigos corretores de erros na mensagem binária.

A primeira característica garante a transparência necessária para a aplicação em *marketing*. O segundo aspecto refere-se à robustez do espalhamento em frequência contra o canal ambiente. As duas últimas qualidades citadas aumentam a probabilidade da marca d'água ser transmitida com sucesso (robustez), quando a mensagem na saída do detector é idêntica à mensagem na entrada do incorporador. A penúltima característica busca reverter os efeitos danosos do canal ambiente para aprimorar a marca d'água detectada, enquanto a última corrige alguns erros ocorridos na extração dos *bits* contidos no sinal marcado recebido.

3.3.2 Análise do Canal Ambiente

No início deste capítulo, definiu-se de forma superficial o canal ambiente como um processo que afeta de várias formas simultaneamente o sinal de áudio transmitido. O primeiro trabalho a tratar da caracterização deste canal foi (STEINEBACH et al., 2002), que foi produzido por alguns dos criadores do serviço de avaliação de métodos de marca d'água *StirMark Benchmark* (PETITCOLAS et al., 2001). Esse estudo identificou, baseado em experimentos práticos, parâmetros da transmissão acústica que estabelecem um modelo para testes automatizados de algoritmos de marca d'água. Esse modelo deve considerar os seguintes componentes:

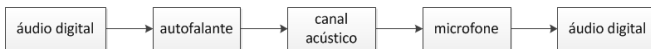
1. **placa de som DA:** produz ruído e distorção de quantização;
2. **amplificação:** modifica a energia do sinal e adiciona ruído devido a amplificadores internos;
3. **alto-falantes:** aplica a resposta em frequência específica do equipamento;
4. **ambiente:** inclui reverberação de sala, ruído ambiente (como conversas ao fundo), e a distância entre os alto-falantes e o microfone aumenta a perda de qualidade de forma proporcional;
5. **microfone:** análogo ao **alto-falante**, impõe a resposta em frequência específica do equipamento;
6. **amplificação do microfone:** semelhante à **amplificação**;
7. **placa de som AD:** semelhante à **placa de som DA**.

Segundo (STEINEBACH et al., 2002), os componentes acima influenciam os sinais transmitidos da seguinte forma:

- adicionando ruído e distorções;
- modificando a amplitude e deslocando a fase;
- quantizando;
- modificando a energia do sinal;
- filtrando, devido à resposta em frequências aplicadas dos *hardwares* e do ambiente.



(a) Completo segundo (STEINEBACH et al., 2002)



(b) Simplificado

Figura 9 – Modelo do canal ambiente

A Figura 9(a) apresenta o modelo descrito acima e permite identificar onde estão os processos mais degradantes do canal ambiente.

Dito isso e em vista da escolha por um algoritmo de espalhamento de frequência, esses processos são reduzidos aos efeitos de filtragem dos alto-falantes e microfone, e a reverberação de sala e ruído ambiente (canal acústico), ilustrado pela Figura 9(b), enquanto os demais efeitos causam danos contra os quais esses algoritmos resistem.

Examinando as filtrações citadas, tem-se dois contextos quanto à resposta ao impulso: plana ou passa-baixas. Na primeira hipótese, o sinal na saída dessa filtragem é uma versão proporcional e atrasada do sinal de entrada, não modificando seu espectro de frequências. Por outro lado, no segundo contexto, as altas frequências da entrada são atenuadas em comparação com as baixas frequências, produzindo um sinal com espectro de frequências afetado em consideração o sinal inicial. Entretanto essas situações não são críticas para o espalhamento de frequência, pois o espalhamento permite considerar o pior caso na incorporação visando suportar este efeito.

Ainda assim, os efeitos do canal acústico são mais severos. Para entendê-los, deve-se analisar o comportamento do som no ar livre desde a fonte emissora (por exemplo, alto-falantes) até o receptor sonoro (como um microfone). A seguir, investiga-se canal acústico e define-se um modelo de simulação do canal ambiente como um todo.

3.3.2.1 Canal Acústico

O som ou a onda sonora é uma onda mecânica de pressão que se propaga de forma longitudinal. Diferente das ondas de luz, o som possui uma condição de existência, que é a necessidade de um meio de propagação. Uma onda sonora pode viajar por diversos meios e materiais, entre os mais comuns estão o ar (meio gasoso), água (meio líquido) e metal (meio sólido) e apresenta características como comprimento de onda, amplitude e frequência.

As ondas sonoras também obedecem aos mesmos fenômenos gerais da propagação ondulatória, sofrendo reflexão, refração, difusão e difração, e fatores como umidade, temperatura e tipo de interação entre as moléculas do meio modificam a velocidade em que estas ondas viajam. Devido a esses fenômenos, o sinal de som emitido por uma fonte sonora não é idêntico ao que chega num receptor. Quando uma onda sonora incide sobre um material, parte da energia é absorvida e outra refletida. Dependendo das características da superfície e da onda incidente, o som refletido é espalhado em direções aleatórias (PASSERO, 2009). As ondas sonoras no interior de uma sala refletem em paredes

e objetos, perdem energia, mudam de direção e sofrem atrasos com a propagação no ar, construindo interferências construtivas e destrutivas entre os sinais sonoros (NEELY; ALLEN, 1979). Um receptor, portanto, recebe o sinal sonoro que percorre o caminho direto sem influência de obstáculos e uma parcela indireta, constituída de todas estas outras ondas que foram refletidas pelo ambiente, que chegam atrasadas e atenuadas. A isto é dado o nome de reverberação, quando a parcela indireta não é distinguida da parcela direta devido o curto intervalo de tempo entre elas. Quando esse intervalo de tempo é suficiente para que haja distinção, esse fenômeno é chamado eco.

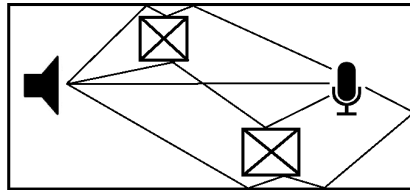


Figura 10 – Ondas refletem no ambiente e chegam atrasadas ao receptor

Basicamente, a reverberação em um ambiente pode ser caracterizada pela resposta ao impulso do ambiente (CARSON; GIESBRECHT; PERRY, 2009). Essa resposta representa um modelo de como o som se comporta em relação àquela exata posição no ambiente e aquele instante de tempo. Variações no ambiente como mudança de lugar de objetos ou do receptor/emissor, obstáculos transitando e até mudanças na umidade e temperatura, podem levar a diferentes respostas ao impulso para um mesmo ambiente (CECCHI et al., 2012).

A característica de multipercurso dos sinais sonoros causa um efeito bem conhecido dos sistemas de telecomunicações, a interferência inter-simbólica (*Inter-Symbol Interference* - ISI). Esta distorção é caracterizada pela interferência de um símbolo transmitido em símbolos seguintes. Acontece em sistemas com multipercurso e canais limitados em banda. Na Figura 11, pode-se ver um exemplo do que esta distorção faz ao sinal. O símbolo enviado é espalhado no tempo de uma maneira que os próximos pulsos que chegam ao receptor são afetados pela contribuição do símbolo anterior, o que pode causar erros na interpretação dos pulsos transmitidos.

Existem algumas maneiras de reduzir os efeitos dessas interferências e uma forma é o aumento do atraso entre os pulsos enviados até que toda a contribuição de um pulso tenha se extinguido. Todavia, essa solução dificilmente é adotada nos dias de hoje, devido à neces-

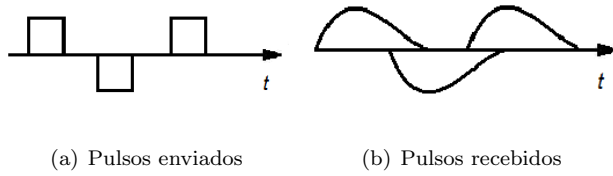


Figura 11 – ISI na transmissão de pulsos

sidade crescente de maiores taxas de transmissão dos sistemas atuais. Outra opção é a equalização do sinal transmitido, que busca inverter o efeito do canal convolutivo através de um filtro e tornar a resposta em frequência do sistema plana.

Além da reverberação, o receptor está sujeito à interferência do ruído ambiente, composto por todo som diferente do emitido, que degrada o som transmitido. Estamos sujeitos ao ruído em qualquer ambiente que estivermos, seja o ruído da ventoinha do computador, o barulho do ar-condicionado, pessoas falando, ou outras fontes.

Portanto, o canal acústico, num dado instante de tempo, pode ser modelado como um filtro linear do tipo FIR somado a um ruído aditivo (PROAKIS, 2001)

$$\hat{\mathbf{s}}_a = \mathbf{s}_a * \mathbf{c}_a + \boldsymbol{\eta}_a, \quad (3.3)$$

em que \mathbf{c}_a é a resposta ao impulso do canal acústico num instante de tempo, \mathbf{s}_a é o áudio emitido, $\boldsymbol{\eta}_a$ representa o ruído acústico e $\hat{\mathbf{s}}_a$ é o áudio resultante após passar pelo canal de transmissão acústico.

3.3.2.2 Canal Ambiente

O canal ambiente apresentado na Figura 9(b) é composto por filtros (alto-falante e microfone) e o canal acústico. Este último foi modelado como um filtro somado a um ruído ambiente, tornando possível definir o canal ambiente, baseado na Eq. (3.3), como

$$\mathbf{r} = (\mathbf{s}_w * \mathbf{c}_{alto} * \mathbf{c}_a + \boldsymbol{\eta}_a) * \mathbf{c}_{mic}, \quad (3.4)$$

sendo \mathbf{r} o sinal marcado recebido, \mathbf{c}_{alto} a resposta ao impulso dos alto-falantes (emissor) e \mathbf{c}_{mic} a resposta ao impulso do canal acústico do

microfone (receptor).

Observando a Eq. (3.4), nota-se que é possível simplificá-la de forma a termos um filtro que compreenda toda transformação que o sinal marcado sofre na transmissão pelo canal acústico somado a um ruído acústico deformado pela resposta do microfone. O filtro citado representa os efeitos combinados dos alto-falantes, do canal acústico e do microfone sobre o sinal \mathbf{s}_w , traduzindo-se e um filtro linear do tipo FIR constante por um período suficientemente curto de tempo. Por outro lado, o ruído acústico é impossível de se modelar, pois tem fontes variáveis. Contudo, aproxima-se o ruído acústico gravado pelo microfone, isto é, o ruído ambiente, como um ruído branco gaussiano filtrado. Dessa forma, o modelo final do canal ambiente é

$$\mathbf{r} = \mathbf{s}_w * \mathbf{c} + \boldsymbol{\eta}_G * \mathbf{c}_n = \hat{\mathbf{s}}_w + \boldsymbol{\eta}_A, \quad (3.5)$$

em que \mathbf{c} é a resposta ao impulso do canal ambiente, $\boldsymbol{\eta}_G$ é o ruído gaussiano branco, \mathbf{c}_n a resposta ao impulso do ruído ambiente, $\boldsymbol{\eta}_A$ o ruído ambiente. A Figura 12 mostra o diagrama de blocos referente à Eq. (3.5).

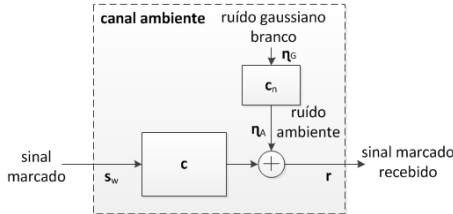


Figura 12 – Modelo final do canal ambiente

As simulações computacionais realizadas neste estudo usam como modelo da Eq. (3.5), cujos filtros \mathbf{c} e \mathbf{c}_n são referentes à uma sala de trabalho (com ruído de conversas, ar condicionado, vento, computador, entre outros) e foram aproximados por testes práticos utilizando alto-falantes e microfones.

Primeiramente, o filtro \mathbf{c}_n é estimado com a gravação do ruído ambiente, sem a emissão de som dos alto-falantes, usando um celular para realizar a gravação. O espectro de potência dessa gravação é estimado e usado no projeto recursivo do filtro FIR, através do método dos mínimos quadrados, que melhor se assemelha ao espectro, estabelecendo o filtro \mathbf{c}_n .

De forma semelhante, o segundo passo utiliza a gravação do ruído

gaussiano branco reproduzido pelos alto-falantes, utilizando o mesmo celular, para calcular a estimativa do espectro de potência do canal ambiente. Entretanto, neste espectro há, além da parcela referente ao filtro \mathbf{c} , a parcela do ruído ambiente. Assim sendo, temos

$$\text{PS}(\hat{\boldsymbol{\eta}}_G) = \text{PS}(\mathbf{c})\text{PS}(\boldsymbol{\eta}_G) + \text{PS}(\boldsymbol{\eta}_A),$$

sendo PS (*Power Spectrum*) a estimativa do espectro de potência e $\text{PS}(\hat{\boldsymbol{\eta}}_G)$, $\text{PS}(\mathbf{c})$, $\text{PS}(\boldsymbol{\eta}_G)$ e $\text{PS}(\boldsymbol{\eta}_A)$ os espectros da gravação do ruído gaussiano branco, do canal ambiente, do ruído gaussiano branco e da gravação do ruído ambiente, respectivamente. Usando o espectro do ruído ambiente obtido anteriormente e estimando o espectro de potência do ruído gaussiano branco, conhecido pelo sistema, chegamos ao espectro de potência do canal \mathbf{c}

$$\text{PS}(\mathbf{c}) = \frac{\text{PS}(\hat{\boldsymbol{\eta}}_G) - \text{PS}(\boldsymbol{\eta}_A)}{\text{PS}(\boldsymbol{\eta}_G)}. \quad (3.6)$$

Mais uma vez, projeta-se o filtro FIR, referente ao canal \mathbf{c} , a partir do espectro de potência definido na Eq. (3.6).

Por fim, com os dois filtros FIR caracterizados, pode-se realizar os testes para o modelo proposto e definir uma forma de combatê-lo. Na Figura 13, observa-se os espectros de potência de um sinal conhecido passando pelo modelo e o mesmo sinal transmitido pelas caixas acústicas e gravado por um microfone em 8 situações diferentes, definidas segundo a Tabela 2. A aproximação é considerada boa e válida para testar a solução, dada a proximidade das duas respostas, definindo assim um modelo de comunicação adequado ao canal ambiente, e os canais pelos quais os testes no Capítulo 6 são feitos.

Dada essa consideração, uma forma de melhorar o sinal recebido, tornando-o mais próximo da saída do incorporador, é equalizar os efeitos do canal ambiente, o que também combate o ISI de acordo com o explicado na Seção 3.3.2.1.

3.4 CONCLUSÃO

Neste capítulo foram apresentados os modelos de comunicação que podem ser aplicados para modelar sistemas de marca d'água, assim como o modelo usado neste trabalho. Foram também discutidos vários algoritmos de marca d'água de áudio, desde a inserção de *bits* nas amostras de áudio até o uso do espalhando em frequência para modular

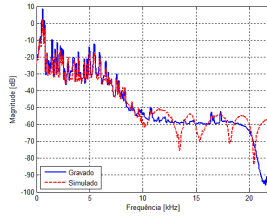
Tabela 2 – Relação das oito posições relativas entre o emissor e o receptor do som

Canal	Emissor/receptor do som	
	Distância	Ângulo
1	30 cm	0°
2	30 cm	90°
3	60 cm	0°
4	60 cm	90°
5	100 cm	0°
6	100 cm	90°
7	150 cm	0°
8	150 cm	90°

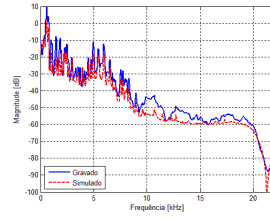
os *bits*.

Tendo em mente que o sistema proposto deve ser robusto aos efeitos do canal ambiente, comentou-se também o estado da arte dos sistemas de marca d'água que tratam dos canais analógico e ambiente. Isso possibilitou a análise do efeito desse canal no áudio, cuja conclusão definiu um modelo para o canal ambiente utilizado neste trabalho.

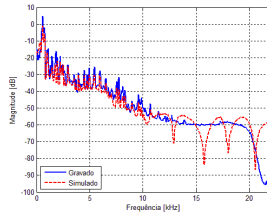
Isto posto, no próximo capítulo, define-se uma marca d'água com alta transparência, detentora de propriedades que possibilitem equalizar os efeitos do canal ambiente.



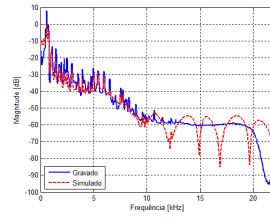
(a) Canal 1



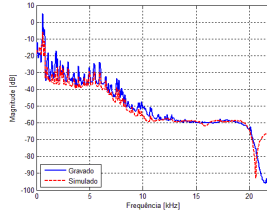
(b) Canal 2



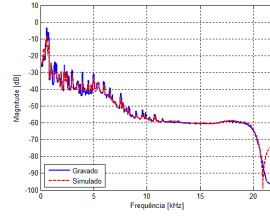
(c) Canal 3



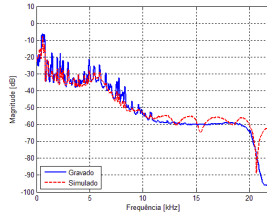
(d) Canal 4



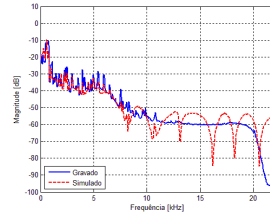
(e) Canal 5



(f) Canal 6



(g) Canal 7



(h) Canal 8

Figura 13 – Espectro de potência do sinal gravado e sinal estimado pelo modelo para diferentes canais

4 SISTEMA PROPOSTO DE MARCA D'ÁGUA DE ÁUDIO

Conforme o exposto no Capítulo 3, os trabalhos que consideram o canal ambiente, conforme apresentado na Tabela 1, são baseados no espalhamento em frequência, exceto (STEINEBACH et al., 2002) que não especifica o tipo de técnica. Dentre esses, somente (WOLFF; BARAS; SICLET, 2010) incorpora a marca no domínio do tempo e executa a equalização do canal ambiente.

A contribuição em (WOLFF; BARAS; SICLET, 2010) se deve ao aumento da robustez da marcação à transmissão no ambiente acústico usando a equalização do canal ambiente. Isto é possível graças a duas etapas: a estimativa do canal ambiente, através da estimativa de canal com redução da interferência (RICE); e sua inversão, produzindo um equalizador forçador a zero (ZFE) não-causal com resposta ao impulso finita. O autor mostra, através de testes com a simulação do canal ambiente, uma redução significativa da BER de 0,2 para 0,0009 quando a taxa de *bits* é de 100 bps.

Todavia, o tempo necessário para que cada estimativa de canal tenha uma boa aproximação do canal real é de aproximadamente 2,3 segundos, um período muito longo no qual nenhuma informação útil é transmitida. Ademais, o autor considera a marcação sempre sincronizada e que o canal é constante durante estimativa e extração dos *bits*, um período de aproximadamente 14 segundos, o que inviabiliza a aplicação prática da solução, uma vez que o canal ambiente causa a dessincronização (XIANG; HUANG, 2006) e é variável (CECCHI et al., 2012).

Assim sendo, a técnica de marca d'água transparente de áudio neste estudo modifica o método introduzido em (WOLFF; BARAS; SICLET, 2010). Primeiramente, para uma aplicação prática, é necessário um método de sincronização robusto ao canal ambiente que não interfira com a marca incorporada e seja transparente. Uma técnica com boas perspectivas foi apresentada em (LIU; INOUE, 2003) e baseia-se em padrões sinusoidais para criar uma marca d'água de espalhamento de frequência. Apesar de (LIU; INOUE, 2003) não comprovarem a robustez de seu processo frente ao canal ambiente, a implementação e testes práticos no desenvolvimento do sistema proposto demonstraram tal atributo. Por último, para o sistema proposto, substituiu-se a forma que a equalização do canal ambiente é feita em (WOLFF; BARAS; SICLET, 2010), objetivando reduzir o período reservado a esse processo e

o tempo de processamento, e permitir imunidade às variações do canal.

A seguir, relata-se as etapas da marca d'água transparente de áudio proposta, ilustradas na Figura 14: incorporador, sincronizador, equalizador e extrator. Ressalta-se que o canal ambiente foi explanado na Seção 3.3.2.2.



Figura 14 – Esquema geral do sistema implementado

4.1 INCORPORADOR

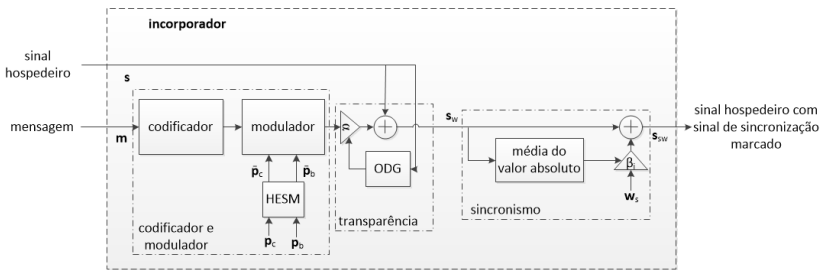


Figura 15 – Diagrama do incorporador do sistema proposto

Na Figura 15, apresenta-se o diagrama da incorporação da marca d'água no sinal de áudio hospedeiro, no qual existem três etapas:

1. **codificador e modulador** - o codificador e o modulador da mensagem m usam, respectivamente, um código corretor de erros e duas sequências pseudoaleatórias binárias (a sequência dos *bits*, $p_b \in [-1, 1]$, e do cabeçalho, $p_c \in [-1, 1]$), produzindo a marca d'água w_a ;
2. **transparência** - adição da marca d'água w_a ao sinal s com garantia de transparência;
3. e **sincronismo** - adição do sinal de sincronização ao sinal hospedeiro marcado, gerando o sinal s_{sw} .

4.1.1 Codificador e Modulador

No codificador, aplica-se algum código corretor de erros ao conjunto de *bits* da sequência \mathbf{m} , criando uma mensagem binária $\mathbf{b} \in [-1, 1]$ contendo \mathbf{m} e *bits* de correção de erro. Esse código visa corrigir erros na etapa de demodulação do extrator. Dentre os muitos tipos de codificadores, como código de repetição, Reed-Solomon, código BCH, código turbo, entre outros, escolheu-se o BCH, devido a sua simplicidade e controle preciso no número de símbolos errados corrigíveis.

Em seguida, a mensagem binária \mathbf{b} é modulada pela sequência pseudoaleatória binária \mathbf{p}_b para formar o sinal \mathbf{w}_r . Como a marca d'água é incorporada de forma contínua, uma seguida da outra ao longo do sinal hospedeiro, usa-se um segunda sequência pseudoaleatória binária, \mathbf{p}_c , como um cabeçalho, indicando o início e o fim de cada mensagem modulada. Desta forma, considerando que as sequências \mathbf{p}_b e \mathbf{p}_c apresentam dimensões $N_b \times 1$ e $N_c \times 1$, respectivamente, o sinal \mathbf{w}_r de comprimento $N = N_c + kN_b$ é determinado por

$$\mathbf{w}_r = [\mathbf{p}_c^t \quad b_0\mathbf{p}_b^t \quad b_1\mathbf{p}_b^t \quad \cdots \quad b_{k-2}\mathbf{p}_b^t \quad b_{k-1}\mathbf{p}_b^t]^t,$$

sendo k o número de *bits* da mensagem \mathbf{b} e b_i o i -ésimo *bit* da mensagem binária \mathbf{b} .

Segundo (WOLFF; BARAS; SICLET, 2010), o sinal \mathbf{w}_r é ajustado seguindo um filtro perceptual derivado do modelo psicoacústico do sinal hospedeiro gerando a marca d'água \mathbf{w}_a . No entanto, os testes mostraram que os melhores resultados são obtidos quando o filtro modelador leva em conta somente o limite de audição humano, desconsiderando o sinal hospedeiro na concepção do filtro modelador, pois a marca continua perceptualmente transparente, porém com mais energia. Neste caso, o filtro pode ser interpretado como um filtro passa altas, pois o formato de sua resposta em frequência é similar ao da resposta de tal filtro, o que é comprovado na Figura 7.

Portanto, para simplificar e reduzir o peso computacional do procedimento, o filtro correspondente ao modelo de sensibilidade do ouvido humano (Figura 7), atua sobre as sequências \mathbf{p}_b e \mathbf{p}_c antes da produção do sinal \mathbf{w}_r , produzindo as sequências $\bar{\mathbf{p}}_b$ e $\bar{\mathbf{p}}_c$, respectivamente. Isto redefine a marca \mathbf{w}_a de forma que esta equivale ao sinal \mathbf{w}_r , em contrapartida a uma versão modelada desse sinal segundo (WOLFF; BARAS; SICLET, 2010), produzindo a seguinte marca d'água, ilustrada na Fi-

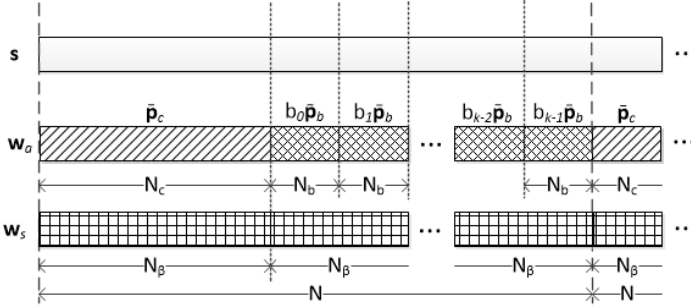


Figura 16 – Esquema gráfico dos sinais incorporados ao sinal hospedeiro

gura 16

$$\mathbf{w}_a = \mathbf{w}_r = [\bar{\mathbf{p}}_c^t \quad b_0 \bar{\mathbf{p}}_b^t \quad b_1 \bar{\mathbf{p}}_b^t \quad \cdots \quad b_{k-2} \bar{\mathbf{p}}_b^t \quad b_{k-1} \bar{\mathbf{p}}_b^t]^t.$$

4.1.2 Transparência

Nesse estágio, a marca \mathbf{w}_a é escalada pelo fator α , balanceando a robustez e a transparência da marca frente o sinal hospedeiro, para, finalmente, somar-se ao sinal hospedeiro \mathbf{s} , originando o sinal hospedeiro marcado \mathbf{s}_w

$$\mathbf{s}_w = \mathbf{s} + \alpha \mathbf{w}_a. \quad (4.1)$$

O cálculo do fator de compromisso, α , da Eq. (4.1) pode ser gerado de diversas formas. Como a imperceptibilidade é mais importante, optou-se por definir α de acordo com o gradiente ODG. Este gradiente possui 5 intervalos que correspondem ao gradiente de diferença subjetiva usada em testes de áudio com humanos (BOSI; GOLDBERG, 2003): $[0, -1)$, imperceptível; $[-1, -2)$, perceptível, mas não incômoda; $[-2, -3)$, levemente incômoda; $[-3, -4)$, incômoda; $[-4, -\infty)$, muito incômoda.

Desta maneira, o valor esperado de ODG seria maior ou igual a -1, contudo, durante a reprodução do áudio marcado no canal ambiente, a degradação e o ruído deste canal mascaram a marca d'água, possibilitando o aumento da energia desta. Por consequência, α pode ser tal que o ODG entre \mathbf{s} e \mathbf{s}_w seja maior ou igual a -2, permitindo que a marca tenha maior energia, contudo mantendo sua transparência ao

ser transmitida pelo canal ambiente.

4.1.3 Sincronismo

Por fim, para a última parte, deve-se criar o sinal de sincronização \mathbf{w}_s de acordo com o descrito em (LIU; INOUE, 2003), para ser somado ao sinal hospedeiro de forma consecutiva ao longo deste, pois \mathbf{w}_s é significamente menor que \mathbf{s}_w . Segundo (LIU; INOUE, 2003), o sinal \mathbf{w}_s é formado pela soma de N_s senoides, com frequências digitais distribuídas de f_0 a $f_0 + N_s - 1$, ponderadas por uma sequência pseudoaleatória binária, $\mathbf{p}_s \in [-1, 1]$, e por uma sequência de ajuste da amplitude das senoides referente ao HESM da Figura 7, **a**. Sendo \mathbf{w}_s de comprimento N , $\sin\left(\frac{2\pi f}{N}\right)$ um sinal de seno de tamanho N e frequência digital f , o sinal de sincronização é definido por

$$\mathbf{w}_s = \sum_{n=0}^{N_s-1} a(n)p_s(n)\sin\left(\frac{2\pi f_0 + n}{N}\right). \quad (4.2)$$

Em seguida, usando o sinal de sincronização retirado da Eq. (4.2), incorpora-se este sinal multiplicado pela média dos valores absolutos do sinal de áudio em intervalos de N_β amostras, denominada β_i , sendo i referente a um intervalo específico, $N = lN_\beta$ e l uma constante inteira positiva. O valor β_i faz a amplitude dos padrões sinusoidais variarem temporalmente com o sinal hospedeiro marcado. Consequentemente, isto pode ser descrito, para um dado intervalo i , por

$$\mathbf{s}_{sw}(i) = \mathbf{s}_w(i) + \beta_i \mathbf{w}_s(i), \quad (4.3)$$

em que $\mathbf{s}_w(i) = [s_w(i), s_w(i-1), s_w(i-2), \dots, s_w(i-N_\beta+1)]^t$, $\mathbf{w}_s(i) = [w_s(i), w_s(i-1), w_s(i-2), \dots, w_s(i-N_\beta+1)]^t$ e $\mathbf{s}_{sw}(i) = [s_{sw}(i), s_{sw}(i-1), s_{sw}(i-2), \dots, s_{sw}(i-N_\beta+1)]^t$ possuem comprimento N_β e representam os elementos dos sinais \mathbf{s}_w a serem modificados, \mathbf{w}_s a serem incorporados, e $\mathbf{s}_{sw}(i)$ resultantes, respectivamente; e β_i é a média dos valores absolutos das amostras de $\mathbf{s}_w(i)$.

4.2 SINCRONIZADOR

O sincronizador é composto pelo detector de sinal de sincronização de (LIU; INOUE, 2003) e pelo ajuste do sinal recebido de acordo

com o diferencial de sincronização detectado, deslocando o sinal recebido pelo número de amostras correspondente ao diferencial.

No detector, a busca pela sincronização é ilustrada na Figura 17, onde (a) são dois sinais de sincronização incorporados, (b) é o sinal de sincronização usado para a busca e (c) é o pico onde o sinal (b) tem o valor máximo da correlação com o sinal marcado em (a). Suponha um intervalo de dessincronização n_d entre os dois inícios dos sinais de sincronização (a) e (b), então há o diferencial $n_t = N - n_d$ do começo do sinal (b) para a posição do pico mostrado em (c).

Dessa forma, a busca da sincronização é dada pelos seguintes passos:

1. o sinal de sincronização (b) é deslocado passo a passo ao longo do sinal recebido \mathbf{r} e a correlação entre este e o sinal (b) é calculada;
2. o pico da correlação calculada, como mostrado em (c), indica a posição de sincronização quando o sinal (b) é deslocado dentro do intervalo de dois sinais de sincronização incorporados;
3. então, o diferencial de dessincronização é calculado $n_t = N - n_d$ e o sinal \mathbf{r} é deslocado de n_t amostras, gerando o sinal sincronizado \mathbf{r}^s .

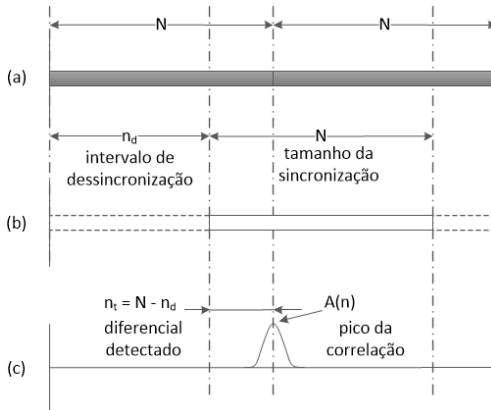


Figura 17 – Diagrama da busca da sincronização pelo sinal de sincronização, sendo (a) dois sinais consecutivos de sincronização incorporados, (b) o sinal de sincronização usado para a busca e (c) o pico onde o sinal (b) tem o valor máximo da correlação com o sinal (a)

4.3 EQUALIZADOR

Em (WOLFF; BARAS; SICLET, 2010), o equalizador é composto pelo estimador do canal acústico e pelo ZFE com resposta ao impulso finita (FIR). O estimador de (WOLFF; BARAS; SICLET, 2010) consiste em um estágio de treinamento baseado na adaptação do algoritmo RICE (CHEN; HUDSON; YAO, 2002), uma vez que este foi desenvolvido especificamente para a estimativa da resposta em frequência de canais acústicos para reverberação de áudio. No entanto, a técnica RICE utiliza um longo período de treinamento, no qual um padrão de treinamento é usado para realizar a estimativa. Esse período é inviável para o objetivo da marca d'água de áudio deste estudo.

Assim sendo, no equalizador proposto, realiza-se a identificação de canal inverso com filtragem adaptativa. A identificação também usa um estágio de treinamento, no qual o canal inverso é estimado de forma adaptativa e contínua, envolvendo uma sequência de treinamento. Em (WOLFF; BARAS; SICLET, 2010), essa sequência foi incorporada no sinal hospedeiro antes da marca d'água, contudo, para a filtragem adaptativa, a sequência pseudoaleatória $\bar{\mathbf{p}}_c$ da marca d'água é utilizada para o treinamento cada vez que o cabeçalho aparece, de forma contínua, aumentando a robustez às variações do canal \mathbf{c} .

Dessa maneira, enquanto em (WOLFF; BARAS; SICLET, 2010) a estimativa do filtro inverso do canal ambiente $\hat{\mathbf{c}}^{-1}$ é feita empregando o canal estimado $\hat{\mathbf{c}}$ na estimativa ótima de mínimos quadrados não-causal, no sistema proposto, a estimativa do filtro inverso é feita diretamente, descartando a etapa de inversão da estimativa do canal ambiente. Finalmente, após estimar o filtro que representa a resposta ao impulso do canal ambiente inverso $\hat{\mathbf{c}}^{-1}$, equaliza-se com esse filtro o sinal \mathbf{r}^s , criando o sinal equalizado $\bar{\mathbf{r}}^s$, visto na Figura 18. A etapa do equalizador é abordada de forma profunda no Capítulo 5.

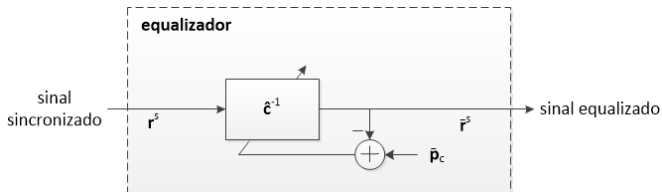


Figura 18 – Diagrama do equalizador do sistema proposto

4.4 EXTRATOR

Por fim, a última etapa possui duas partes: o demodulação e o decodificador. A primeira parte extrai os *bits* modulados pela sequência pseudoaleatória na etapa de codificação e modulação do incorporador. Em seguida, a segunda fase corrige eventuais erros nos *bits* extraídos, diminuindo a probabilidade erro da extração da mensagem transmitida.

4.4.1 Demodulação

Para a demodulação, o sinal equalizado é filtrado pelo filtro de Wiener IIR não-causal, que visa minimizar o erro médio quadrático entre o sinal $\bar{\mathbf{p}}_b$ e o sinal equalizado. Esse filtro é calculado primeiramente dividindo a densidade espectral de potência (*Power Spectrum Density* - PSD) do sinal de referência $\bar{\mathbf{p}}_b$, previamente calculada, pela PSD do segmento correspondente ao sinal de referência no sinal equalizado $\bar{\mathbf{r}}^s$ com N_b amostras, denominado $\bar{\mathbf{r}}_b$. O resultado desta divisão é então transformado para o domínio do tempo com a transformada inversa de Fourier (IFFT) e o filtro de Wiener IIR não-causal é a parte real desta transformação. Desta forma, o filtro de Wiener correspondente ao sinal de referência é

$$\mathbf{w}_b = \text{real} \left(\text{IFFT} \left(\frac{\text{PSD}(\bar{\mathbf{p}}_b)}{\text{PSD}(\bar{\mathbf{r}}_b)} \right) \right).$$

O filtro de Wiener da sequência $\bar{\mathbf{p}}_b$ almeja preparar o sinal para a demodulação dos *bits* inseridos, reduzindo a influência dos demais sinais presentes (o sinal hospedeiro recebido, o sinal de sincronização recebido e o ruído ambiente). Após a filtragem, o sinal \mathbf{v}_b obtido é correlacionado com $\bar{\mathbf{p}}_b$, resultando em um *bit* da mensagem estimada $\hat{\mathbf{b}}$

$$\hat{b}_j = \begin{cases} 1 & , \text{ se } \frac{1}{N_b} \sum_{i=1}^{N_b} v_b(i) \bar{p}_b(i) > 0 \\ 0 & , \text{ se } \frac{1}{N_b} \sum_{i=1}^{N_b} v_b(i) \bar{p}_b(i) \leq 0 \end{cases}.$$

4.4.2 Decodificação

A mensagem estimada é, enfim, corrigida, caso necessário, usando o decodificador para o código BCH utilizado no incorporador, produzindo a saída do extrator a mensagem extraída $\hat{\mathbf{m}}$. Todo esse processo

pode ser visto na Figura 19.

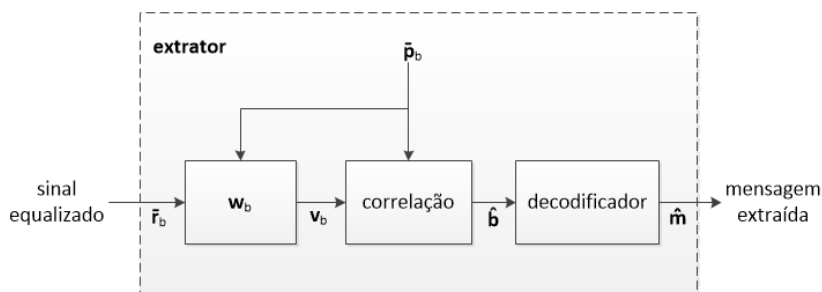


Figura 19 – Diagrama do extrator do sistema proposto

4.5 CONCLUSÃO

Neste capítulo foram expostos os componentes que formam o sistema proposto de marca d'água transparente em áudio. Foram abordados a incorporação, a sincronização, a equalização e a extração da marca, sendo estas são as etapas que compõe a marca proposta. No incorporador, explicou-se como a marca é incorporada no sinal hospedeiro e como sua transparência é alcançada, enquanto no sincronizador e no equalizador, as modificações do canal ambiente são revertidas. Por fim, o extrator não propõe nenhuma novidade frente ao existente na literatura. Porém, o equalizador será melhor explicado no próximo capítulo, onde a técnica existente na literatura é comparada com a proposta e as vantagens e limitações do proposto são discutidas.

5 EQUALIZADOR ADAPTATIVO

A equalização de canais não ideais de transmissão está presente na literatura de forma consistente (MIYOSHI; KANEDA, 1988), inclusive combinada com marcas d'água (JUNIOR; MIRANDA, 2003; WOLFF; BARAS; SICLET, 2010). Esse processo visa reverter os efeitos indesejados causados pelo canal, tais como atenuações e amplificações em diferentes frequências que modificam o sinal transmitido, descaracterizando-o muitas vezes e causando problemas como a ISI.

O objetivo de um equalizador é tornar plana a resposta em frequência de um sinal afetado por um canal específico. Idealmente, um sinal equalizado é idêntico ao sinal original transmitido, isto significa que a convolução do canal $c(n)$ com o equalizador $c^{-1}(n)$ deve resultar em um impulso deslocado no tempo em D amostras (DINIZ, 2002)

$$c(n) * c^{-1}(n) = \delta(n - D). \quad (5.1)$$

Um equalizador conhecido é o forçador a zero (ZFE), cuja natureza é eliminar completamente os efeitos do canal de transmissão forçando o erro introduzido pelo ISI a zero. O ZFE desconsidera o ruído aditivo da transmissão. Portanto, se o canal apresentar nulos espectrais, o equalizador possuirá enormes ganhos nas faixas de frequência dos nulos. Esse ganho, por sua vez, pode amplificar o ruído aditivo nas componentes de frequência correspondentes aos nulos espectrais (PROAKIS, 2001; DINIZ, 2002). Outra origem deste problema está no fato do canal ambiente, principalmente o canal ambiente de uma sala, ser muitas vezes de fase não-mínima (NEELY; ALLEN, 1979). Consequentemente, a inversa deste canal pode ser causal e instável ou não-causal e estável (NAYLOR; GAUBITCH, 2010), possuindo diversas técnicas de cálculo, como o método dos mínimos quadrados (*Least Squares* - LS) ou o teorema inverso de múltiplas entradas/saídas (*Multiple-Input/Output Inverse Theorem* - MINT) (MIYOSHI; KANEDA, 1988). Devido às características de um sistema de marca d'água, na qual há somente uma entrada e uma saída, a técnica MINT não é aplicável, assim, (WOLFF; BARAS; SICLET, 2010) faz uso do LS descrito em (MIYOSHI; KANEDA, 1988).

Em vista dessa possível distorção na presença de ruído, outro tipo de equalização é bastante aplicada: equalizador por minimização do erro quadrático médio (*Minimum Mean Square Error Equalizer* - MMSEE) entre um sinal de treinamento e o sinal filtrado pelo equali-

zador. Nesse caso, o equalizador consiste na estimativa de uma função de transferência para compensar as distorções lineares causadas pelo canal. Então, esta solução é um compromisso entre a minimização da ISI e redução do ruído, dado que o equalizador ótimo da Eq. (5.1) nunca é alcançado. Por isso, para o canal ambiente, a função de transferência do equalizador não é exatamente a inversa do canal convolutivo (DINIZ, 2002).

Mesmo buscando evitar os problemas devido ao ISI e ao ruído aditivo, ambos equalizadores ZFE e MMSEE são suscetíveis quando essas distorções são severas (PROAKIS, 2001). Nestes casos, o equalizador não-linear por decisão realimentada (*Decision-Feedback Equalizer* - DFE) pode ser uma alternativa. Combinando a solução do MMSEE com a realimentação através de um segundo filtro adaptativo, o DFE usa a decisão de símbolos anteriores para anular a interferência do ISI e do ruído na decisão do símbolo atual (DINIZ, 2002). Apesar de potencialmente reduzir os problemas de equalização do canal ambiente, esta abordagem sofre de um viés, a realimentação inerente deste equalizador pode causar o acúmulo de erros, caso a saída do equalizador esteja incorreta, isto é, o símbolo decodificado não está correto. Isto provoca picos de erros que podem comprometer toda a extração da marca d'água.

Embora o DFE possa ser, teoricamente, mais robusto aos efeitos do canal ambiente, optou-se por empregar o MMSEE neste trabalho em virtude da fácil comparação entre o MMSEE e o ZFE. Desta forma, enquanto (WOLFF; BARAS; SICLET, 2010) adota um ZFE não-adaptativo, propõe-se utilizar o MMSEE adaptativo.

Considerando o canal ambiente variante no tempo para grande intervalos de tempo (maiores que 20 ms) e invariante no tempo para curtos períodos (até aproximadamente 20 ms), a estimativa do equalizador proposto é feita em intervalos próximos a 20 ms e aprimorada periodicamente a cada segmento do sinal gravado correspondentes à sequência de cabeçalho, buscando adequar o equalizador às mudanças do canal ambiente.

Assim, o equalizador proposto é o MMSEE adaptativo e não-cego, sendo possível, através das estatísticas do sinal emitido e recebido, avaliar os efeitos do canal ambiente e revertê-los. O equalizador apresentado em (WOLFF; BARAS; SICLET, 2010) é o ZFE não-adaptativo e não-cego utilizando inversor LS não-causal, entretanto sua sequência de treinamento é independente da marca d'água inserida, representando uma perda de capacidade durante o treinamento.

A seguir, os conceitos de filtragem adaptativa e os filtros usados

serão introduzidos. Em seguida, a estimativa do canal ambiente inverso é descrita, onde também são comentadas suas limitações. Por fim, a equalização propriamente dita é explicada.

5.1 FILTRAGEM ADAPTATIVA

A ferramenta de filtragem adaptativa geralmente trata de situações sem definição clara ou que variam ao longo do tempo. Esses filtros são considerados variantes por ter a atualização de seus parâmetros no decorrer do tempo para atingir determinado requerimento. A variação dos coeficientes depende do sinal de entrada, de um sinal de referência e do algoritmo de minimização do erro adotado.

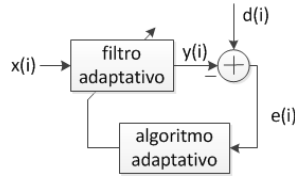


Figura 20 – Configuração básica de um filtro adaptativo

A Figura 20 apresenta a estrutura básica de um filtro adaptativo, em que $x(i)$ é o sinal de entrada, $d(i)$ é o sinal desejado, $e(i)$ é o erro definido como a diferença entre a saída do filtro $y(i)$ e o sinal de referência $d(i)$, e i é a i ésima iteração. O algoritmo de atualização dos coeficientes usa esse erro para melhorar o desempenho do filtro, minimizando esse erro ao quadrado, pois a busca da minimização do erro quadrático é uma função convexa dos coeficientes do filtro.

Segundo (DINIZ, 2002), a especificação de um sistema adaptativo descrito na Figura 20 consiste em três itens:

1. Aplicação - O tipo de aplicação depende dos sinais medidos do ambiente e escolhidos pelo projetista para serem os sinais de entrada e referência. Há muitas aplicações as quais este tipo de solução é viável, entre elas está o cancelamento de eco, identificação de sistemas, conformação de feixe adaptativo, equalização e cancelamento de ruído.
2. Estrutura do filtro adaptativo - A estrutura do filtro influencia na complexidade computacional do processo e no número necessário de iterações para chegar a um determinado nível de desempenho

esperado. Basicamente há duas maneiras de se implementar um filtro digital adaptativo, e o que as distingue é a forma da resposta ao impulso. Há a resposta ao impulso finita e a resposta ao impulso infinita. A saída de um filtro FIR é simplesmente a soma ponderada das últimas amostras de entrada. A ordem do filtro FIR indica sua memória. Enquanto o filtro IIR faz o mesmo procedimento, porém adiciona a soma ponderada de amostras de saídas passadas. Portanto, o filtro FIR é considerado não recursivo e o IIR recursivo. (A recursividade do IIR pode, muitas vezes, levar à instabilidade devido ao crescimento das amostras de saída que por sua vez fazem os pesos do filtro crescerem exponencialmente, levando ao *overflow*.)

3. Algoritmo - Esse o processo que influencia o desempenho que se quer obter na utilização do filtro. O algoritmo é determinado pela definição do método de busca ou algoritmo de minimização, função objetivo e natureza do sinal de erro.

O filtro adaptativo tem como objetivo encontrar o conjunto de coeficientes $\theta(i)$ que minimize a função objetivo $F(x(i), y(i), d(i))$. Os coeficientes $\theta(i)$ devem convergir para o filtro ótimo $\theta_0(i)$ que produz o erro quadrático mínimo, idealmente zero, aproximando $y(i)$ de $d(i)$. Os métodos mais difundidos na literatura são: método de Newton, métodos Quasi-Newton e método do gradiente (DINIZ, 2002). O primeiro busca o mínimo de uma aproximação de segunda ordem da função custo através da atualização de parâmetros iterativa. Os métodos Quasi-Newton são a simplificação do método de Newton, os quais utilizam a recursividade para realizar a minimização. Por último, o método do gradiente adota a direção oposta ao vetor gradiente da função objetivo para encontrar o ponto de mínimo desta função.

Três métodos diferentes serão abordados neste trabalho, sendo dois derivados do método do gradiente e um de Newton, são eles: mínimo quadrado médio (LMS), mínimo quadrado médio normalizado (NLMS) e mínimo quadrado recursivo (RLS). Todos os algoritmos buscam da solução do filtro de Wiener e, assim sendo, esta solução e os três algoritmos são explicados em seguida.

5.1.1 Filtro de Wiener

Uma das funções de custo mais utilizadas é a do erro médio quadrático. Seguindo o exemplo da Figura 20, onde o erro é dado por

$$e(i) = d(i) - y(i),$$

, podemos defini-la por

$$F[e(i)] = \varepsilon(i) = E\{e^2(i)\} = E\{d^2(i) - 2d(i)y(i) + y^2(i)\}. \quad (5.2)$$

Considerando o filtro com resposta ao impulso finita, temos que

$$y(i) = \sum_{j=0}^{N_\theta-1} \theta_j(i)x(i-j) = \boldsymbol{\theta}^t(i)\mathbf{x}(i),$$

sendo $\mathbf{x}(i) = [x(i), x(i-1), \dots, x(i-N_\theta+1)]^t$ o vetor de entrada e $\boldsymbol{\theta}(i) = [\theta_0(i), \theta_1(i), \dots, \theta_{N_\theta-1}(i)]^t$ os coeficientes do filtro e N_θ a ordem do filtro. Em um caso linear com um filtro de coeficientes fixos e em um ambiente estacionário, pode-se reescrever a Eq. (5.2) desta forma

$$\begin{aligned} \varepsilon(i) &= E\{d^2(i)\} - 2\boldsymbol{\theta}^t(i)E\{d(i)\mathbf{x}(i)\} + \boldsymbol{\theta}^t(i)E\{\mathbf{x}(i)\mathbf{x}^t(i)\}\boldsymbol{\theta}(i) \\ &= E\{d^2(i)\} - 2\boldsymbol{\theta}^t(i)\mathbf{P} + \boldsymbol{\theta}^t\mathbf{R}\boldsymbol{\theta}(i), \end{aligned}$$

em que \mathbf{R} é a matriz de auto correlação do vetor de entrada e \mathbf{P} é o vetor de correlação cruzada entre o sinal de referência e o sinal de entrada. O vetor gradiente do MSE relacionado aos coeficientes do filtro é

$$\mathbf{g}_\theta(i) = \frac{\nabla \varepsilon(i)}{\nabla \boldsymbol{\theta}(i)} = -2\mathbf{P} + 2\mathbf{R}\boldsymbol{\theta}(i). \quad (5.3)$$

Se igualarmos o vetor gradiente a zero e considerarmos \mathbf{R} não singular, os valores ótimos dos coeficientes do filtro de Wiener são dados por

$$\boldsymbol{\theta}_0 = \mathbf{R}^{-1}\mathbf{P}. \quad (5.4)$$

Infelizmente, na prática não conhecemos com precisão \mathbf{R} e \mathbf{P} e as estimativas dessa matriz e desse vetor não são exatamente práticas. Assim, os algoritmos adaptativos buscam alcançar uma aproximação da solução de Wiener sem realizar essas estimativas.

5.1.2 LMS

O algoritmo de busca LMS os coeficientes de um filtro adaptativo capaz de minimizar o erro médio quadrático. Para isso, são utilizadas simplificações do método do gradiente. O LMS é um dos métodos mais utilizados devido a sua baixa complexidade computacional, sua convergência em ambiente estacionário e convergência, na média, para a solução de Wiener da Eq. (5.4).

Se considerarmos o método do gradiente, que atualiza os coeficientes seguindo o sentido contrário do gradiente multiplicado por um passo de convergência, e estimativas das matrizes \mathbf{R} e \mathbf{P} , denominadas $\hat{\mathbf{R}}$ e $\hat{\mathbf{P}}$, temos:

$$\begin{aligned}\boldsymbol{\theta}(i+1) &= \boldsymbol{\theta}(i) - \mu \hat{\mathbf{g}}_{\theta}(i) \\ &= \boldsymbol{\theta}(i) + 2\mu(\hat{\mathbf{P}}(i) - \hat{\mathbf{R}}(i)\boldsymbol{\theta}(i))\end{aligned}\quad (5.5)$$

em que $\hat{\mathbf{g}}_{\theta}(i)$ representa uma estimativa do gradiente da função objetivo em relação aos coeficientes do filtro. Estimativas instantâneas para $\hat{\mathbf{R}}$ e $\hat{\mathbf{P}}$ são

$$\begin{aligned}\hat{\mathbf{R}}(i) &= \mathbf{x}(i)\mathbf{x}^t(i) \\ \hat{\mathbf{P}}(i) &= d(i)\mathbf{x}(i).\end{aligned}$$

Pode-se assim definir a estimativa $\hat{\mathbf{g}}_{\theta}(i)$ como

$$\begin{aligned}\hat{\mathbf{g}}_{\theta}(i) &= -2d(i)\mathbf{x}(i) + 2\mathbf{x}(i)\mathbf{x}^t(i)\boldsymbol{\theta}(i) \\ &= 2\mathbf{x}(i)(-d(i) + \mathbf{x}^t(i)\boldsymbol{\theta}(i)) \\ &= -2e(i)\mathbf{x}(i).\end{aligned}$$

Substitui-se $\hat{\mathbf{g}}_{\theta}(i)$ na Eq. (5.5) e obtém-se

$$\boldsymbol{\theta}(i+1) = \boldsymbol{\theta}(i) + 2\mu e(i)\mathbf{x}(i). \quad (5.6)$$

A Equação (5.6) representa a equação de atualização dos coeficientes do filtro para o algoritmo LMS. O fator μ é um passo de atualização que deve ser escolhido cuidadosamente, pois um valor muito pequeno pode fazer o algoritmo demorar a convergir, enquanto um valor grande pode fazer os coeficientes convergirem de forma rápida, porém com um erro em excesso grande, ou pode instabilizar o algoritmo.

Devido às simplificações e estimações assumidas para a realização do método, o algoritmo nunca irá chegar à solução ótima de Wiener.

Todavia, os coeficientes do filtro tendem à solução de Wiener com um erro mínimo.

5.1.3 NLMS

Muitos algoritmos surgiram a partir de derivações do LMS. O NLMS ou LMS normalizado é um desses algoritmos. Variações de potência em $x(i)$ levam a variações significativas no ajuste dos coeficientes de $\theta(i)$ no LMS. Na tentativa de melhorar o comportamento de convergência do algoritmo, buscou-se modificar o fator de atualização μ de forma que este varie de acordo com o sinal de entrada $x(i)$, minimizando o erro instantâneo (DINIZ, 2002).

Faz-se a normalização de μ em relação à potência de $x(i)$, deixando a função de atualização do filtro assim

$$\theta(i+1) = \theta(i) + \frac{\mu}{\epsilon + \mathbf{x}^t(i)\mathbf{x}(i)} e(i)\mathbf{x}(i).$$

O parâmetro ϵ é um fator de segurança para não haver divisão por zero. E μ é introduzido para controlar desajustes, pois todas as derivações para a construção deste algoritmo estão baseadas em valores instantâneos dos erros quadráticos e não no erro quadrático médio (DINIZ, 2002).

5.1.4 RLS

Uma vez que os algoritmos oriundos do LMS almejam a minimização da soma dos quadrados das diferenças entre o sinal desejado e a saída do filtro adaptativo, se as amostras do sinal de entrada são recebidas a cada iteração, esse problema pode ser calculado recursivamente. O resultado disto são os algoritmos RLS (DINIZ, 2002). Esses algoritmos são conhecidos por buscar uma rápida convergência mesmo quando o espalhamento dos autovalores da matriz de correlação do sinal de entrada é grande. Por esta razão, este algoritmo possui um desempenho excelente para sistemas variantes no tempo. Contudo, estas vantagens advêm ao custo do aumento da complexidade computacional e de alguns problemas de estabilidade.

Resumidamente, o algoritmo RLS pode ser interpretado como uma aproximação estocástica do algoritmo de Newton ou como a solução exata de um problema de estimativa usando mínimos quadrados.

Dessa maneira, sabendo que o método de Newton usa uma aproximação quadrática da superfície da função custo, expande-se o erro médio quadrático por série de Taylor

$$\begin{aligned}\varepsilon(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\boldsymbol{\theta}(i)} &= \varepsilon[\boldsymbol{\theta}(i)] + \nabla\varepsilon|_{\boldsymbol{\theta}=\boldsymbol{\theta}(i)}^t [\boldsymbol{\theta} - \boldsymbol{\theta}(i)] \\ &\quad + \frac{1}{2}[\boldsymbol{\theta} - \boldsymbol{\theta}(i)]^t \mathbf{H}[\boldsymbol{\theta}(i)] [\boldsymbol{\theta} - \boldsymbol{\theta}(i)],\end{aligned}\quad (5.7)$$

sendo $\mathbf{H}[\boldsymbol{\theta}(i)] = \frac{\partial^2 \varepsilon(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}(i)}$ a hessiana de $\varepsilon(\boldsymbol{\theta})$. Para determinar a função de atualização do filtro, diferencia-se a Eq. (5.7) em relação a $\boldsymbol{\theta}$ e iguala-se a zero em $\boldsymbol{\theta} = \boldsymbol{\theta}(i)$. Desta forma,

$$\nabla\varepsilon[\boldsymbol{\theta}(i+1)] = \nabla\varepsilon[\boldsymbol{\theta}(i)] + \frac{1}{2}2\mathbf{H}[\boldsymbol{\theta}(i)][\boldsymbol{\theta}(i+1) - \boldsymbol{\theta}(i)] = 0,$$

e, assim,

$$\boldsymbol{\theta}(i+1) = \boldsymbol{\theta}(i) - \mathbf{H}^{-1}[\boldsymbol{\theta}(i)]\nabla\varepsilon[\boldsymbol{\theta}(i)]. \quad (5.8)$$

Considerando a Eq. (5.3) e a hessiana $\mathbf{H}^{-1}[\boldsymbol{\theta}(i)] = (2\mathbf{R})^{-1} = \frac{1}{2}\mathbf{R}^{-1}$, tem-se que a Eq. (5.8) é expandida para

$$\begin{aligned}\boldsymbol{\theta}(i+1) &= \boldsymbol{\theta}(i) - \frac{1}{2}\mathbf{R}^{-1}[-2\mathbf{P} + 2\mathbf{R}\boldsymbol{\theta}(i)] \\ &= \mathbf{R}^{-1}\mathbf{P} = \boldsymbol{\theta}_0.\end{aligned}\quad (5.9)$$

Como a superfície é quadrática, o método de Newton converge em uma única iteração, a partir de qualquer ponto inicial, desde que a inversão de \mathbf{R} seja perfeita. Assim, devido a um possível mal condicionamento de $\mathbf{H}[\boldsymbol{\theta}(i)] = 2\mathbf{R}$, usa-se $\hat{\mathbf{H}}[\boldsymbol{\theta}(i)] = 2\mathbf{R} + 2\epsilon\mathbf{I}$, em que ϵ é um valor que garanta que a hessiana não seja nula.

Então, considera-se o algoritmo de Newton com regularização

$$\boldsymbol{\theta}(i+1) = \boldsymbol{\theta}(i) - \mu[\epsilon\mathbf{I} + \mathbf{R}]^{-1}[\mathbf{P} - \mathbf{R}\boldsymbol{\theta}(i)],$$

em que o algoritmo RLS melhora a aproximação de $\epsilon\mathbf{I} + \mathbf{R}$ usando todos os vetores de entrada passados, introduzindo memória variável no sistema, para reduzir a importância dos dados mais antigos na atualização do filtro, e aplica aproximações instantâneas. Essas modificações são

descritas pelas seguintes aproximações

$$\begin{aligned}\hat{\mathbf{R}} &= \frac{1}{i+1} \sum_{j=0}^i \lambda^{i-j} \mathbf{x}(j) \mathbf{x}^t(j) \\ \mu(i) &= \frac{1}{i+1} \\ \epsilon(i) &= \lambda^{i-j} \frac{\epsilon}{i+1} \\ \mathbf{P} - \mathbf{R}\boldsymbol{\theta}(i) &\simeq \mathbf{x}(i)[d(i) - \mathbf{x}^t(i)\boldsymbol{\theta}(i)] = \mathbf{x}(i)e(i),\end{aligned}$$

sendo $0 \ll \lambda < 1$ o fator de esquecimento do algoritmo, $\mu(n)$ o passo de adaptação que tende para zero quando $n \rightarrow \infty$ e $\epsilon(n)$ a regularização que desaparece para $n \rightarrow \infty$ e com $\epsilon > 0$. Com estas alterações, tem-se a adaptação dos coeficientes do filtro adaptativo pelo algoritmo RLS

$$\begin{aligned}\boldsymbol{\theta}(i+1) &= \boldsymbol{\theta}(i) + \frac{1}{i+1} \left[\frac{\lambda^{i+1}}{i+1} \epsilon \mathbf{I} + \frac{1}{i+1} \sum_{j=0}^i \lambda^{i-j} \mathbf{x}(j) \mathbf{x}^t(j) \right]^{-1} \mathbf{x}(i)e(i) \\ &= \boldsymbol{\theta}(i) + \left[\lambda^{i+1} \epsilon \mathbf{I} + \sum_{j=0}^i \lambda^{i-j} \mathbf{x}(j) \mathbf{x}^t(j) \right]^{-1} \mathbf{x}(i)e(i).\end{aligned}\quad (5.10)$$

5.2 ESTIMATIVA DO CANAL AMBIENTE INVERSO

Em concordância com o comentado no Capítulo 4, a estimativa do canal ambiente inverso é feita usando a estrutura de equalização de canal. Essa estrutura é ilustrada na Figura 21, utilizando a mesma notação da Seção 5.1.

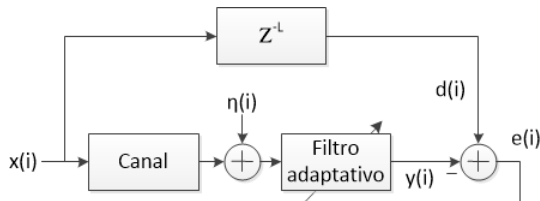


Figura 21 – Diagrama de equalização de canal

A estrutura adaptada ao sistema de marcação digital de áudio é demonstrada na Figura 22. Conforme discutido anteriormente, usa-se

um sinal de treinamento para estimar o canal ambiente inverso, denominado sinal desejado na Figura 20. Esse sinal é, neste trabalho, a sequência de cabeçalho e , para que a estimativa seja correta, o sinal de entrada é uma versão modificada dessa sequência. Dessa forma, o sinal de entrada é o segmento do sinal sincronizado correspondente à \bar{p}_c .

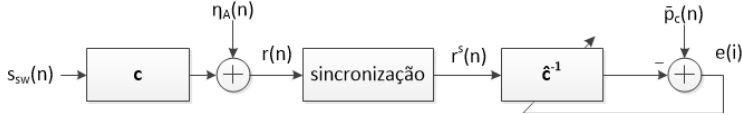


Figura 22 – Diagrama de identificação de canal inverso com filtragem adaptativa baseado na marca d'água

Observando a Figura 22 e aplicando a Eq. (3.5), pode-se concluir que o sinal de entrada neste diagrama é dado por

$$\begin{aligned}
 r^s(n) &= \hat{s}_{sw}(n) + \eta_A(n) \\
 &= \hat{s}(n) + \hat{p}_c(n) + \hat{w}_s(n) + \eta_A(n) \\
 &= c(n) * s_{sw}(n) + \eta_A(n) \\
 &= c(n) * s(n) + c(n) * \alpha \bar{p}_c(n) + c(n) * \beta w_s(n) + \eta_A(n), \quad (5.11)
 \end{aligned}$$

em que a sincronização não é demonstrada neste equacionamento, mas considerada perfeita e $\eta_A(n)$ é o ruído ambiente. Portanto, o sinal de erro é definido como

$$\begin{aligned}
 e(n) &= \bar{p}_c(n) - \hat{c}^{-1}(n) * r^s(n) \\
 &= \bar{p}_c(n) - \hat{c}^{-1}(n) * c(n) * s_{sw}(n) - \hat{c}^{-1}(n) * \eta_A(n) \\
 &= \bar{p}_c(n) - \hat{c}^{-1}(n) * c(n) * \alpha \bar{p}_c(n) - \hat{c}^{-1}(n) * c(n) * s(n) \\
 &\quad - \hat{c}^{-1}(n) * c(n) * w_s(n) - \hat{c}^{-1}(n) * \eta_A(n). \quad (5.12)
 \end{aligned}$$

A partir da Figura 22 e da Eq. (5.12), podemos analisar que o filtro adaptativo \hat{c}^{-1} é calculado buscando minimizar o sinal de erro, isto é, a diferença entre $\bar{p}_c(n)$ e $r^s(n)$.

Contudo, o cálculo desse filtro restringe os sinais envolvidos a localmente estacionários durante o período de estimação. Uma vez que a aproximação do canal ambiente a um filtro constante é aceitável para esse período e sinais de áudio são estocásticos e de média nula para o mesmo período, de aproximadamente 20 ms, o esquema proposto produz resultados verossímeis. Ao final da adaptação desse filtro, a conclusão é a estimativa do canal ambiente inverso, que é usado como

equalizador do restante do sinal recebido antes da extração da mensagem.

5.3 EQUALIZAÇÃO DO SINAL RECEBIDO

De posse do filtro correspondente à estimativa do canal ambiente inverso, para que a mensagem possa ser extraída com sucesso, filtra-se o restante do sinal recebido sincronizado até a próxima sequência de cabeçalho. O resultado é um sinal que se aproxima do sinal hospedeiro com sinal de sincronização marcado, pois elimina parte da influência do canal ambiente.

Na prática, o canal ambiente nunca pode ser anulado com o equalizador MMSE, pois o ruído ambiente interfere na estimativa e no erro médio quadrático final, impedindo este de tender a zero no infinito. Contudo, o equalizador adaptativo do sistema proposto de marca d'água de áudio contribui para melhorar a extração da mensagem ao mesmo tempo que aumenta a verossimilhança entre o sinal equalizado e o sinal hospedeiro com sinal de sincronização marcado.

5.4 CONCLUSÃO

Neste capítulo foram discutidos os aspectos envolvendo a equalização do canal ambiente, desde o tipo de equalização escolhida, o MMSE adaptativo, com estimativa adaptativa do canal ambiente inverso usando os métodos LMS, NLMS e RLS.

Por fim, todo o sistema de marca d'água digital em áudio proposto foi explicado e em seguida os testes referentes a esse sistema são discutidos.

6 RESULTADOS

Os resultados obtidos serão demonstrados de duas maneiras distintas, a fim de uma abordagem mais completa dos mesmos. Sendo assim, vemos tanto a sua forma básica (sem equalização), quanto com o auxílio da equalização com estimação do canal inverso e com o canal inverso ótimo (usado na simulação), para canais simulados e testes reais (aqui, exceto o canal inverso ótimo). Com ambas as interpretações dos resultados, podemos obter uma consciência real das contribuições deste trabalho e verificar os possíveis pontos a desenvolver.

Nos métodos de comparação presentes na literatura que avaliam e comparam sistemas de marca d'água, normalmente leva-se em conta a taxa de erro de *bit* (BER) (KRAETZER; DITTMANN; LANG, 2006). Essa taxa representa o percentual dos bits transmitidos que são corretamente decodificados após a etapa de extração, que aqui foi utilizada como parâmetro de comparação neste trabalho.

Foram utilizados, para teste, 50 áudios, com frequência de amostragem 44100 Hz, correspondentes a diversas gravações divididas em 5 gêneros musicais: blues, eletrônico, jazz, pop e rock. Os testes foram realizados utilizando uma mensagem com 16 *bits* pseudoaleatórios de informação incorporados 25 vezes repetidas a cada áudio com o uso do codificador BCH e extraíndo estas 25 repetições 5 vezes, num total de 2000 *bits* por áudio e totalizando 20000 *bits* extraídos para cada gênero. Mais uma condição importante a ser comentada é o fato da sincronização ser aplicada em todos os testes, garantindo que os resultados são referentes a todo sistema proposto, e não somente à parte de equalização.

O primeiro teste demonstra o ganho proporcionado pelo uso do codificador BCH frente à extração sem o corretor de erros BCH. Desta maneira, foram usadas as condições descritas acima, exceto pela presença ou não do BCH em cada parte do teste, simulando os canais citados no Capítulo 3. O segundo teste é aplicado para apresentar o comportamento da marca d'água proposta frente os 8 canais simulados e os mesmo áudios comentados. O terceiro teste é semelhante ao segundo, no entanto é feito em todas as posições especificadas na Tabela 2 em uma sala de escritório real, com caixas de som de um computador como fonte emissora e com o microfone de um celular como receptor, isto é, as mesmas condições usadas para criar os 8 canais simulados na Seção 3.3.2.2.

Todos os testes foram feitos usando as seguintes definições dos

elementos básicos do sistema proposto: $N_c = 1024$, $N_b = 256$, $k = 64$, $f_0 = 500$ Hz, $N_s = 15$, $N_\beta = 1024$ e uma taxa de *bits* úteis de aproximadamente 40 bps. A seguir estes testes são descritos com maiores detalhes e seus resultados discutidos.

6.1 TESTE 1 - COMPARAÇÃO ENTRE O USO OU NÃO DO BCH

Códigos corretores de erros binários, usados em conjunto com marcas d'água, são extensamente estudados na literatura. Contudo, a adição desses códigos causa a redução da taxa de *bits* efetivos transmitidos e o aumento da capacidade computacional necessária, dado que os corretores usam parte dos *bits* transmitidos para corrigir a informação binária útil usando técnicas matemáticas e estatísticas que consomem tempo de cálculo. Sendo assim, optou-se por verificar a eficácia do uso do código corretor de erros BCH e, por conseguinte, a necessidade do uso desse.

Usando os canais simulados apresentados na Capítulo 3, realizou-se a extração, sem equalização, dos 2000 *bits* de cada um dos 50 áudios usando a codificação BCH e novamente todo esse processo sem a codificação BCH. Para uma comparação correta, os 16 *bits* de informação são codificados com a BCH(63,16) ou são combinados com 47 *bits* zeros em seguida, gerando 63 *bits* codificados que são incorporados e extraídos dos áudios. O resultado disso é apresentado em dois gráficos, a fim de favorecer a visualização e interpretação dos resultados, mostrados na Figura 23.

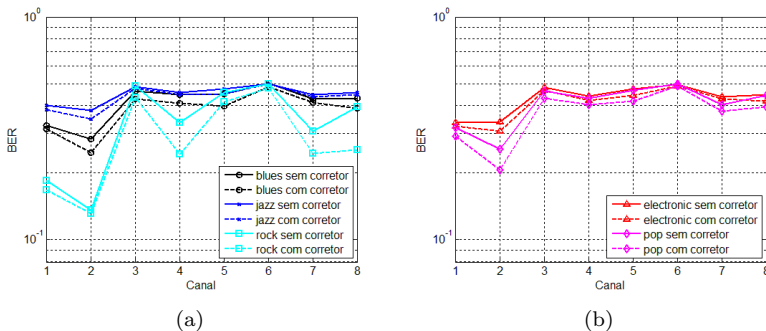


Figura 23 – Comparação entre o uso ou não da codificação BCH

A partir desses resultados, nota-se a redução da BER em praticamente todos os canais e gêneros musicais. Ainda que essa redução, em alguns pontos, mostre-se pequena ou bastante próxima a zero, nota-se que de maneira alguma é indiferente. Com isso em mente e ainda considerando a baixa complexidade computacional necessária para usar esse código corretor de erros e o fato do canal ambiente ser um canal tão adverso, decidiu-se pelo uso da codificação BCH para tornar o sistema proposto ainda mais robusto.

6.2 TESTE 2 - COMPORTAMENTO DO SISTEMA PROPOSTO EM CANAIS SIMULADOS

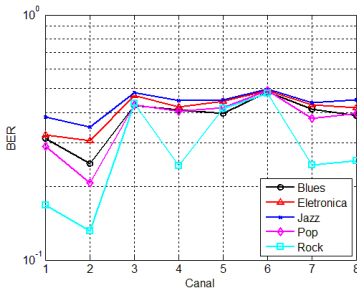
Com a decisão pelo uso do código corretor de erros BCH feita na seção anterior, compara-se o comportamento do sistema proposto em 5 casos diferentes: sem equalização, com equalização usando o método LMS, o NLMS, o RLS e, por fim, usando o canal inverso ideal. Devido às diferenças de escalas, os resultados foram divididos em 6 gráficos, sendo um com os resultados sem equalização de todos os gêneros e os 5 demais com os resultados agrupados de cada gênero separadamente, conforme apresentado na Figura 24.

A Figura 24(a) exibe a limitação da marca d'água muito frágil frente o canal ambiente simulado, visto que poucos gêneros atingem uma BER inferior a 0,15 (um número muito insatisfatório, pois representa uma média de 15 *bits* errados a cada 100 *bits* extraídos) e, mesmo assim, apenas para alguns dos canais. Sendo assim, é entendida a necessidade de trabalhar o sinal recebido para que seja possível extrair a informação de forma confiável.

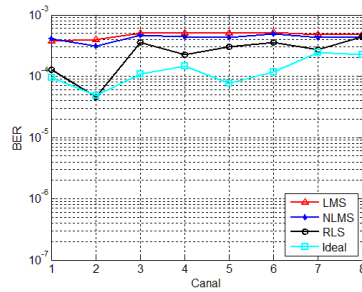
Nas demais figuras, pode-se perceber três principais fatos em uma rápida observação:

1. com a equalização, o erro foi reduzido para abaixo de 10^{-3} ;
2. os melhores resultados foram obtidos usando o RLS, próximo do ideal em diversos canais;
3. e, ainda que usando o canal ambiente inverso ideal, o erro dificilmente torna-se inferior a 10^{-5} .

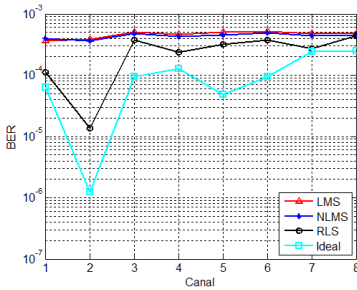
Dessa maneira, observa-se que a redução na BER com o uso da equalização, e o surgimento de menores erros devido ao uso da RLS, comportando-se conforme esperado. Porém os erros mínimos do canal inverso ideal expressos nas curvas das Figura 24 foram absolutamente inesperados.



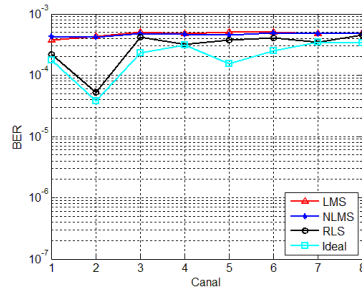
(a) Sem equalização



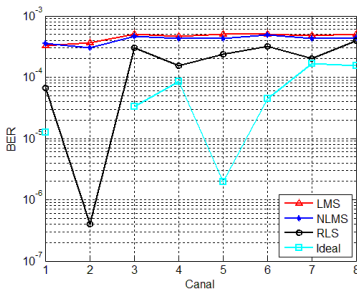
(b) Blues



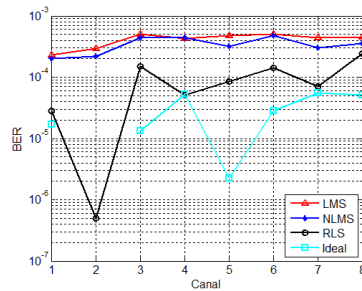
(c) Eletrônico



(d) Jazz



(e) Pop



(f) Rock

Figura 24 – Comportamento do sistema proposto frente canais simulados

Os altos valores de erro mínimo devem-se às interferências introduzidas pelos demais sinais componentes do sinal recebido. Na análise

feita na Seção 5.2, a Equação (5.12) expõe os elementos que são combinados para formar o sinal de erro da estimativa do canal inverso. Consequentemente, mesmo que o canal inverso $\hat{c}^{-1}(n)$ fosse ideal, o sinal de erro seria constituído pelos elementos provenientes dos sinais hospedeiro $s(n)$, de sincronização $w_s(n)$ e de ruído ambiente $\eta_A(n)$. Finalmente, isso torna impossível atingir a BER nula.

No tocante às curvas, o esperado era que fossem mais homogêneas, de maneira que erros aumentassem conforme a posição de gravação era afastada da fonte sonora. Contudo, o resultado mostrou que o conjunto formado pelo receptor (celular), a fonte emissora (alto-falantes) e a sala onde as gravações foram feitas atua de forma impactante no sinal sonoro, dessa maneira, percebe-se que em alguns momentos, mesmo a maiores distâncias, a qualidade do sinal recebido é melhor, do que quando em alguns pontos à distâncias menores. Isso ocorre devido a mudança de ângulo referente tanto à fonte emissora, quanto aos obstáculos na sala e ainda em decorrência da alteração da característica de reverberação da sala quando relacionada a angulação e posicionamento do receptor. Em decorrência desses fatos, fica evidente a necessidade de um pré-processamento no sinal recebido para a extração da informação.

Realizando uma análise segregada por gêneros musicais, percebe-se que alguns gêneros possuem melhores resultados que outros. Os menores erros foram encontrados nos testes para os gêneros pop e rock, seguidos por eletrônico e, por fim, blues e jazz. Os espectros de potência de cada um dos trechos dos áudios usados nos testes são representados individualmente por cada curva presente na Figura 25.

Ao examinar esses gráficos, constata-se que os três gêneros com os piores resultados apresentam, em ao menos um áudio, baixas magnitudes (< -40 dB) principalmente, mas não só, nas faixas de frequência mais audíveis (20 Hz a 10 kHz), causadas pelas baixas energias dos sinais. Essa baixa energia torna a marca d'água inserida no áudio ainda mais fraca (com menos energia) para que a transparência continue nos limites permitidos, contribuindo para o efeito negativo na BER. Por consequência, os dois piores gêneros têm muitos áudios com baixa energia, enquanto o eletrônico que só possui um áudio abaixo dessas magnitudes consegue resultados aceitáveis. Por outro lado, tanto os áudios de pop quanto de rock tem maiores energias nestas frequências, permitindo maior robustez da marca d'água.

Por fim, estes fatos permitem compreender que a energia do sinal hospedeiro tem um grande peso na robustez da marca d'água incorporada através do sistema proposto neste trabalho. Resumidamente,

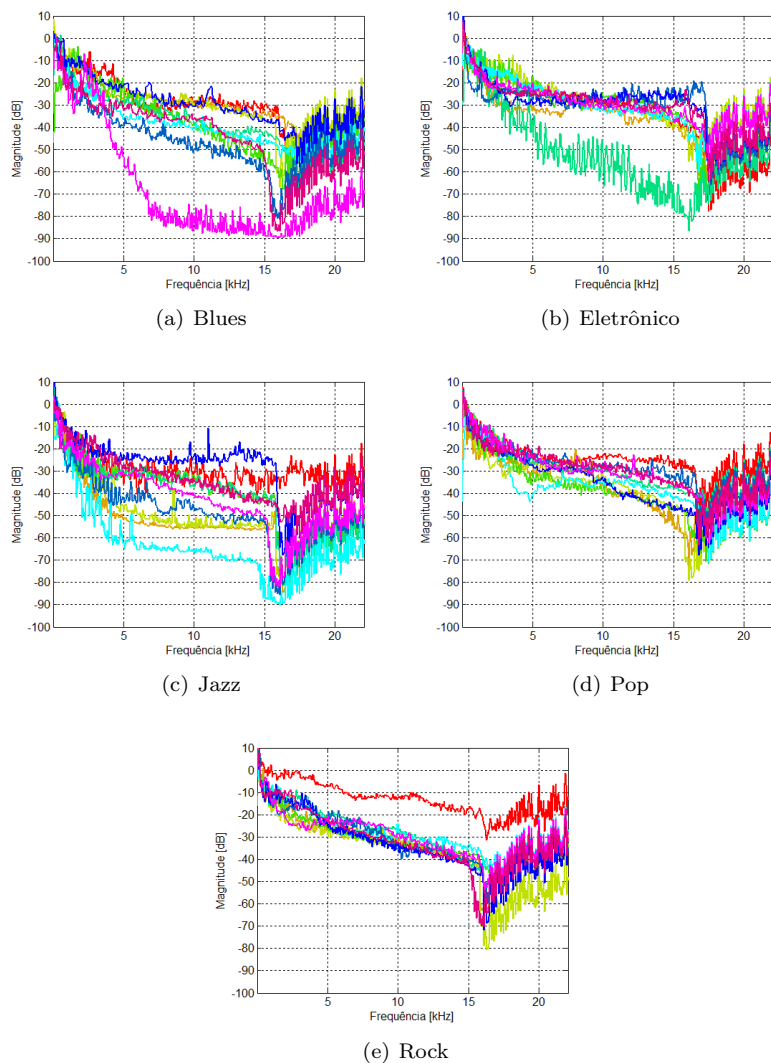


Figura 25 – Espectro de potência dos trechos dos áudios usados nos testes

áudios com maiores energias, como os encontrados nos gêneros pop e rock, oferecem melhores condições e, assim, menores erros na extração

da informação inserida.

Tabela 3 – BER média dos 8 canais

Gênero	Equalizador				
	Nenhum	LMS	NLMS	RLS	Canal inverso
Blues	0,38496	0,00047	0,00042	0,00026	0,00013
Eletrônico	0,41301	0,00046	0,00044	0,00026	0,00011
Jazz	0,43713	0,00047	0,00046	0,00032	0,00023
Pop	0,37679	0,00045	0,00041	0,00021	0,00006
Rock	0,29600	0,00042	0,00034	0,00009	0,00003

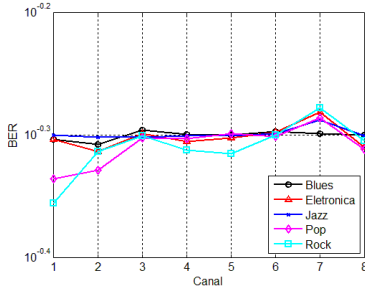
Sabendo que em (WOLFF; BARAS; SICLET, 2010) foi alcançada uma BER inferior a 10^{-4} para a taxa de 40 bps e considerando a média para os 8 canais apresentada na Tabela 3, os valores atingidos pelo sistema proposto com o método RLS e o canal inverso ideal são comparáveis ao obtido em (WOLFF; BARAS; SICLET, 2010). Apesar de superiores, os erros nesses casos levam em consideração mais canais e suas sincronizações são feitas durante a extração ao contrário das sincronizações perfeitas do trabalho anterior. Dessa forma, conseguiu-se alcançar erros similares entre este trabalho e (WOLFF; BARAS; SICLET, 2010) com o uso do sincronismo e da equalização propostos.

6.3 TESTE 3 - COMPORTAMENTO DO SISTEMA PROPOSTO EM SITUAÇÃO REAL (SALA DE ESCRITÓRIO)

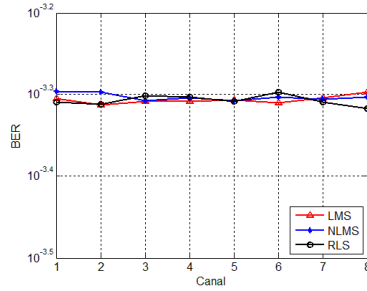
Considerando-se este o teste prático, seus resultados terão a responsabilidade de apontar a eficácia do sistema de marca d'água proposto em um dos cenários para o qual esse sistema foi projetado. Nesse cenário, reproduziu-se o arranjo entre a fonte emissora e o receptor de maneira idêntica ao utilizado para extrair os canais simulados apresentados na Seção 3.3.2.2. As BERs obtidas da extração dos áudios gravados em cada um dos arranjos citados são exibidos na Figura 26.

A exigência da equalização para permitir a extração com confiança dos *bits* de informação inseridos se torna perceptível ao verificar a BER alcançada sem o uso da equalização na Figura 26(a). O erro encontrado sem equalização fica em torno de 0,5 ($10^{-0,3}$), o que significa em um sistema binário o erro máximo, ou seja, a total incerteza.

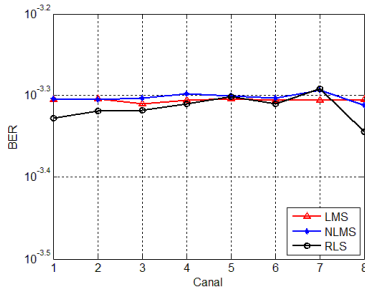
Visto isso, o uso da equalização já reduziu esse erro para valores



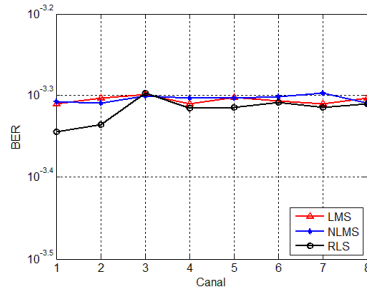
(a) Sem equalização



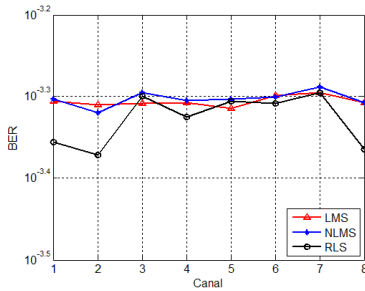
(b) Blues



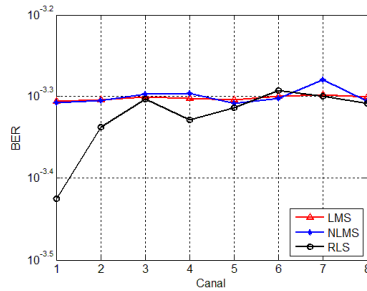
(c) Eletrônico



(d) Jazz



(e) Pop



(f) Rock

Figura 26 – Comportamento do sistema proposto em situação real (sala de escritório)

mais razoáveis, próximos a $5 \cdot 10^{-4}$ ($10^{-3,3}$), visíveis nas Figuras 26(b) a 26(f). Algumas das conclusões encontradas na seção anterior estão

presentes neste teste em diferentes proporções, como a equalização também reduziu o erro, os melhores resultados foram obtidos usando o RLS, as menores e maiores BERs dos gêneros seguem a mesma dinâmica do teste anterior, e fatores independentes ao sinal emitido, como ângulo relativo entre fonte e receptor e obstáculos da sala, interferem de forma imprevisível na extração.

Essas conclusões apenas comprovam, conforme esperado, e sem quaisquer contradição, o deduzido no Teste 2. Consequentemente, entende-se que:

- a equalização reduz drasticamente o erro na extração da informação incorporada;
- o método de convergência mais rápida, o RLS, produz a melhor aproximação do canal ambiente inverso e, portanto, a maior confiabilidade média dos *bits* extraídos;
- os gêneros com maiores energias permitem maior robustez à marca e, por isso, reduzem a BER;
- o canal ambiente é fortemente degradante e interfere significativamente no sistema de marca d'água;
- e, finalmente, o sistema proposto provou ser uma alternativa válida, todavia, incipiente, para contornar as características rigorosas do canal ambiente.

6.4 CONCLUSÃO

Observando-se a comparação previamente explanada entre a extração sem equalização da informação inserida com o uso de um código corretor de erros ou não. Pode-se perceber que o código corretor de erros utilizado, o BCH(63,16), reduziu a BER quando aplicado e, dada sua baixa complexidade, decidiu-se por incorporá-lo ao sistema proposto.

Já numa análise referente aos métodos de adaptação da estimativa do canal ambiente inverso para diferentes canais, primeiramente simulados e depois para reais, observa-se que sinais hospedeiros com maior energia tornam a marca mais robusta, pois permitem a maior energia do sinal incorporados dentro do limite de transparência. Por fim, o melhor resultado deste teste se comportou como o esperado, compatível com o método adaptativo de maior velocidade de convergência,

o RLS, nos levando a uma ultima percepção da eficiência do sistema proposto.

Ao final do capítulo foram apresentados os resultados do sistema em uma situação real, utilizando um celular e caixa de som como receptor e fonte emissora, respectivamente, nas mesmas posições do teste anterior. Extraiu-se as mesmas conclusões do teste anterior, ainda que os valores das BERs não tenham sido equivalentes. Dessa forma, constatou-se que a equalização reduz o erro; o método RLS produz os melhores resultados; os áudios com maior energia permitem maior robustez; o canal ambiente é fortemente degradante; e que o sistema proposto é uma alternativa viável para a aplicação proposta.

7 CONCLUSÕES E PERSPECTIVAS

Este estudo apresentou um novo sistema de marca d'água digital em áudio fosse altamente transparente com objetivo de aplicabilidade em *marketing*.

Inicialmente, foram expostas e discutidas as principais características e propriedades das marcas d'água digitais, relacionando-as com aplicações reais. Dentre as características gerais das marcas d'água digitais, duas tem papel fundamental para o desenvolvimento do estudo aqui desenvolvido. Sendo a transparência a de maior relevância, uma vez que a aplicabilidade em *marketing* exige que o som original não seja prejudicado pela marca d'água, afetando-o de forma perceptiva. Ainda nessa linha, outra característica vital para a aplicabilidade aqui sugerida é a robustez ao canal ambiente, definida como a capacidade da marca d'água de sobreviver às interferências geradas pelo canal de transmissão, de modo a manter a integridade da marca, para que ela cumpra seu objetivo.

Verificou-se, então, a interpretação da marca d'água digital por um modelo de comunicação, de maneira que dentro os modelos existentes, apresentaram-se os de comunicação tradicional, básico e informado no transmissor. Discutiu-se também os métodos tradicionais de marcas d'água em áudio e o estado da arte dos sistemas de marca d'água em áudio específicos ao canal analógico e ambiente ou robustos a esses. Observou-se que para a aplicação, o sistema proposto deveria ter as seguintes propriedades: usar algum modelo psicoacústico do HAS; utilizar espalhamento em frequência; reduzir as ações negativas do canal; e aplicar códigos corretores de erros na mensagem binária.

As técnicas de grande relevância e amplo espectro na literatura direcionadas a aplicações em áudio, como codificação dos *bits* menos significativos, ocultamento de eco, espalhamento em frequência, entre outras, foram descritas por meio de discussões, diagramas e ilustrações. Por fim, a análise das últimas metodologias que tratam da resistência ao canal ambiente permite compreender e delinear soluções para o problema da dessincronização e das características ruidosas e degradantes do canal ambiente em situações reais, problema esse observado em estudos prévios que tratam de temas semelhantes e em testes prévios a este trabalho.

Diferenciando-se de sua maioria que trata de sistemas de marca d'água em áudio e utilizam somente uma forma de marcação em seus sistemas, o sistema aqui apresentado usa duas marcações diferentes,

uma para sincronia e outra para carregar a informação. Decidiu-se pelo uso de duas marcas com funções diferentes por dois motivos: garantir a transparência e permitir o pré-processamento. Existem técnicas que permitem a autossincronização, porém as exigências de transparência e robustez impedem seu uso.

Dessa forma, a separação em duas marcas distintas possibilitou que ambas desempenhem seus papéis com eficiência sem perder a transparência desejada. Contudo, a ideia de usar duas marcas foi imposta pelo objetivo de modificar a etapa de pré-processamento (equalização do canal) do sistema base deste trabalho. Assim, a marca d'água usada para a equalização é similar o bastante a marca utilizada no sistema base para ser possível comparar os resultados de ambas.

O sistema proposto apresenta o uso do limite de audição humano e um ganho adaptativo para garantir a transparência necessária, equivalente ao gradiente ODG maior que -2. Também mostra a aplicação do sinal de sincronização para sincronizar o sinal recebido de forma a possibilitar a extração da marca após os efeitos de dessincronização do canal ambiente. E, por último, a utilização da extração da informação, incorporada da mesma maneira que em estudos prévios, para identificar os *bits* transmitidos e de um código corretor de erros para corrigir os erros nos *bits* extraídos.

Embora na literatura existam referências a equalização do sinal recebido durante o pré-processamento anterior a extração efetiva da informação, a técnica adaptativa de estimar o canal ambiente inverso diretamente é inovação, uma vez que, em trabalhos anteriores, a equalização é feita estimando primeiramente o canal ambiente e, em seguida, invertendo-o com o inversor LS não-causal. Essa técnica de inversão minimiza o problema da inversão do canal ambiente, muitas vezes de fase não-mínima, gerando um filtro não-causal. Conquanto isto seja uma solução elegante, há perda de informação no canal produzido e a possibilidade de elevação do erro causado pelo ganho do ruído em algumas frequência reduz o destaque desta técnica.

Em vista disso, a equalização adotando estimativa de canal ambiente inverso com filtragem adaptativa produz o filtro necessário sem os problemas descritos acima na técnica de estudos prévios. Essa estimativa, descrita no Capítulo 5, encontra de maneira iterativa um filtro que combate as ações do canal a partir de uma segmento da marca d'água. Porém, ainda há problemas, como a interferência dos demais sinais presentes no sinal recebido que aumentam o erro na estimativa do canal ambiente e a chance da necessidade de mais iterações para alcançar uma estimativa adequada.

Todos os resultados apresentado no Capítulo 6 indicam que o sistema proposto se mostrou efetivo para canais ambiente, sejam eles simulados segundo o modelo de canal ambiente composto por um filtro linear e ruído aditivo ou reais. A princípio, demonstrou-se a possibilidade de aumentar a robustez do sistema utilizando o código corretor de erros BCH(63,16). Mesmo que a redução no erro não seja expressiva, a baixa complexidade computacional e o fato de qualquer ganho em uma situação tão adversa ser importante contribuíram para a decisão pelo uso desse código corretor de erros no sistema proposto.

Com a definição do uso do BCH, foram realizados testes com diferentes canais simulados e diversos áudios, comprovando-se as melhorias possíveis que a equalização proporciona à extração da marca. Porém concluiu-se que a estimativa do canal inverso não alcança os resultados ótimos, representados pelos resultados da extração aplicando a equalização com o canal ambiente inverso. Além disso, o método RLS resultou nos menores erros, como esperado, e os áudios dos gêneros com maiores energias, como pop e rock, também produziram as menores BERs. Por fim, em uma comparação com os resultados apresentados em trabalhos anteriores, corrobora-se a eficácia do sistema proposto em um cenário mais complexo que no método anterior, pois seus resultados são similares.

Já os testes reais mostraram que a sincronização proposta é confiável, porém sem a equalização este sistema muito transparente torna-se impraticável. Por outro lado, guardadas as proporções, o sistema proposto se comportou de forma similar frente os canais simulados, com o RLS e os gêneros com maior energia extraindo a informação com maior confiança e comprovando a hipótese de que o canal ambiente seria o maior obstáculo no uso de marcas d'água na aplicação em *marketing*. Por fim, o sistema proposto provou ser uma alternativa válida, todavia, incipiente, para contornar as características rigorosas do canal ambiente.

Isso posto, futuros trabalhos para encontrar melhores maneiras de equalizar o canal ambiente, usando por exemplo o equalizador por decisão realimentada, podendo melhorar as taxas de extração da marca. Outro exemplo é a ampliação dos cenários de testes reais e simulados, utilizando outros aparelhos tanto para a fonte emissora quanto para o receptor, como televisões, microfones, diferentes caixas de som, rádios, diferentes celulares, *tablet*, outros computadores, entre outras fontes; e diferentes ambientes, como salas de estar, salas de recepção, quartos, e restaurantes; com o objetivo de aumentar a variabilidade dos testes e encontrar as melhores definições para o sistema proposto se tornar um

produto viável.

Referências

- ARNOLD, M.; SCHMUCKER, M.; WOLTHUSEN, S. *Techniques and Applications of Digital Watermarking and Content Protection*. [S.l.]: Artech House, 2003. (Artech House computer security series).
- BARAS, C.; MOREAU, N.; DYMARSKI, P. Controlling the inaudibility and maximizing the robustness in an audio annotation watermarking system. *Audio, Speech, and Language Processing, IEEE Transactions on*, v. 14, n. 5, p. 1772–1782, 9 2006.
- BENDER, W. et al. Techniques for data hiding. *IBM Syst. J.*, IBM Corp., Riverton, NJ, USA, v. 35, n. 3-4, p. 313–336, 9 1996.
- BONEY, L.; TEWFIK, A. H.; HAMDY, K. N. Digital watermarks for audio signals. In: *International Conference on Multimedia Computing and Systems*. Hiroshima: [s.n.], 1996. p. 473–480.
- BOSI, M.; GOLDBERG, R. *Introduction to Digital Audio Coding and Standards*. [S.l.]: Kluwer Academic, 2003.
- CARSON, M.; GIESBRECHT, H.; PERRY, T. *Surround Sound Impulse Response*. July 2009.
- CECCHI, S. et al. An adaptive multiple position room response equalizer in warped domain. In: *20th European Signal Processing Conference*. [S.l.: s.n.], 2012.
- CHEN, J. C.; HUDSON, R. E.; YAO, K. Fast frequency-domain acoustic channel estimation with interference cancellation. In: *ICASSP*. [S.l.]: IEEE, 2002. p. 1709–1712.
- COSTA, M. Writing on dirty paper. *IEEE Transactions on Information Theory*, v. 29, n. 3, p. 439–441, 1983.
- COX, I. et al. *Digital Watermarking and Steganography*. 2. ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2008.
- COX, I.; MILLER, M. L.; MCKELLIPS, A. L. Watermarking as Communications With Side Information. In: *Proceedings International Conference Multimedia Computing Systems*. [S.l.: s.n.], 1999. p. 1127–1141.

CVEJIC, N. *Algorithms for Audio Watermarking and Steganography*. Dissertao (Academic Dissertation) — University of Oulu, 6 2004.

DINIZ, P. *Adaptive Filtering: Algorithms and Practical Implementation*. [S.l.]: Kluwer Academic Pub, 2002.

DUTTA, P.; BHATTACHARYYA, D.; KIM, T.-H. Data Hiding in Audio Signal: A Review. *International Journal of Database Theory and Application*, v. 2, 2009.

EGGERS, J.; GIROD, B. *Informed Watermarking*. [S.l.]: Kluwer Academic Publishers, 2002. (Kluwer international series in engineering and computer science).

FASTL, H.; ZWICKER, E. *Psychoacoustics*. [S.l.]: Springer, 2007. (Springer series in information sciences).

GRUHL, D.; BENDER, W.; LU, A. Echo Hiding. In: *Information Hiding*. Berlin, Germany: Springer, 1996. v. 1174, p. 293–315.

HAITSMA, J. et al. Audio watermarking for monitoring and copy protection. In: *Proceedings of the 2000 ACM workshops on Multimedia*. New York, NY, USA: ACM, 2000. (MULTIMEDIA '00), p. 119–122.

HUANG, C.-H. et al. Information delivery systems for car passengers without networking capabilities. In: *Proceedings of the multimedia 8th Pacific Rim conference on Advances in multimedia information processing*. Berlin, Heidelberg: Springer-Verlag, 2007. (PCM'07), p. 539–548.

HUANG, J.; WANG, Y.; SHI, Y. A blind audio watermarking algorithm with self-synchronization. In: *Circuits and Systems, 2002. ISCAS 2002. IEEE International Symposium on*. [S.l.: s.n.], 2002. v. 3, p. 627–630.

JUNIOR, A. N. L.; MIRANDA, M. D. Técnicas de equalização adaptativa usadas em sistemas de comunicação digital. In: MACKENZIE (Ed.). *Caderno de Pós-Graduação em Engenharia Elétrica*. [S.l.: s.n.], 2003.

KANG, X.; YANG, R.; HUANG, J. Geometric Invariant Audio Watermarking Based on an LCM Feature. *IEEE Transactions on Multimedia*, v. 13, n. 2, p. 181–190, 4 2011.

- KIM, H. et al. Audio watermarking techniques. *Intelligent Watermarking Techniques*, p. 185–218, 2004.
- KRAETZER, C.; DITTMANN, J.; LANG, A. Transparency benchmarking on audio watermarks and steganography. p. 60721L–60721L–12, 2006.
- LAZIC, N.; AARABI, P. Communication over an acoustic channel using data hiding techniques. *Multimedia, IEEE Transactions on*, v. 8, n. 5, p. 918–924, 10 2006.
- LIE, W.-N.; CHANG, L.-C. Robust and high-quality time-domain audio watermarking based on low-frequency amplitude modification. *Multimedia, IEEE Transactions on*, v. 8, n. 1, p. 46–59, 2 2006.
- LIU, Z.; INOUE, A. Audio watermarking techniques using sinusoidal patterns based on pseudorandom sequences. *IEEE Trans. Cir. and Sys. for Video Technol.*, IEEE Press, Piscataway, NJ, USA, v. 13, n. 8, p. 801–812, 8 2003.
- MIYOSHI, M.; KANEDA, Y. Inverse filtering of room acoustics. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, v. 36, n. 2, p. 145–152, 1988.
- NAKAYAMA, A. et al. Digital watermarks for audio signal based on psychoacoustic masking model. *Electron. Comm. Jpn. Pt. III*, v. 86, p. 65–75, 12 2003.
- NAYLOR, P.; GAUBITCH, N. *Speech Dereverberation*. [S.l.]: Springer London, 2010. (Signals and Communication Technology).
- NEELY, S. T.; ALLEN, J. B. Invertibility of a room impulse response. *Journal of the Acoustical Society of America, ASA*, v. 66, n. 1, p. 165–169, 1979.
- OMOLOGO, M.; SVAIZER, P. Acoustic source location in noisy and reverberant environment using csp analysis. In: *Acoustics, Speech, and Signal Processing, 1996. ICASSP-96. Conference Proceedings., 1996 IEEE International Conference on*. [S.l.: s.n.], 1996. v. 2, p. 921–924.
- PAINTER, T.; SPANIAS, A. Perceptual Coding of Digital Audio. In: *in Proceedings of the IEEE*. [S.l.: s.n.], 2000.
- PASSERO, C. R. M. *Avaliação e adequação acústica de escritórios panorâmicos através de medições e simulações*. Tese (Doutorado) — Universidade Federal do Paraná, 2009.

PETITCOLAS, F. A. P. et al. Public automated web-based evaluation service for watermarking schemes: Stirmark benchmark. p. 575–584, 2001.

PROAKIS, J. *Digital Communications*. [S.l.]: McGraw-Hill, 2001. (McGraw-Hill Series in Electrical and Computer Engineering).

QIAN, K. et al. Audio annotation watermarking with robustness against DA/AD conversion. In: QIAN, K. et al. (Ed.). *Media Forensics and Security*. [S.l.]: SPIE, 2010. (SPIE Proceedings).

SHANNON, C. E. Channels with side information at the transmitter. *IBM J. Res. Dev.*, IBM Corp., Riverton, NJ, USA, v. 2, n. 4, p. 289–293, 10 1958.

SHIN, S. et al. A robust audio watermarking algorithm using pitch scaling. In: *Digital Signal Processing, 2002. DSP 2002. 2002 14th International Conference on*. [S.l.: s.n.], 2002. v. 2, p. 701 – 704.

STEINEBACH, M. et al. Audio watermarking quality evaluation: Robustness to da/ad processes. In: *Proceedings of the International Conference on Information Technology: Coding and Computing*. Washington, DC, USA: IEEE Computer Society, 2002. (ITCC '02), p. 100–.

TACHIBANA, R. Audio Watermarking for Live Performance. In: *Proc. of SPIE Int. Conf. on Security and Watermarking of Multimedia Contents V*. Santa Clara, USA: [s.n.], 2003. v. 5020, p. 32–43.

TACHIBANA, R. et al. An audio watermarking method using a two-dimensional pseudo-random array. In: . [S.l.: s.n.], 2002. v. 82, n. 10, p. 1455–1469.

TERHARDT, E. Calculating virtual pitch. *Hearing Research*, v. 1, n. 2, p. 155–182, 3 1979.

WOLFF, E.; BARAS, C.; SICLET, C. Toward robustness of audio watermarking systems to acoustic channels. In: *Proceedings of EUSIPCO*. Aalborg, Danemark: [s.n.], 2010. p. 4. Département Images et Signal Département Images et Signal.

WU, C.-P.; SU, P.-C.; KUO, C.-C. J. Robust and efficient digital audio watermarking using audio content analysis. *Multimedia Syst.*, Springer-Verlag New York, Inc., Secaucus, NJ, USA, v. 8, n. 5, p. 382–392, 12 2000.

XIANG, S.; HUANG, J. Analysis of Quantization-based Audio Watermarking to D/A and A/D Conversions. *I. J. Network Security*, v. 3, n. 3, p. 230–238, 2006.

XIANG, S.; HUANG, J. Robust Audio Watermarking Against the D/A and A/D conversions. *CoRR*, 2007.

XIANG, Y. et al. Effective Pseudonoise Sequence and Decoding Function for Imperceptibility and Robustness Enhancement in Time-Spread Echo-Based Audio Watermarking. *IEEE Transactions on Multimedia*, v. 13, n. 1, p. 2–13, 2 2011.

